

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2005, 04:38:43 ; Search time 18292 Seconds  
(without alignments)  
11663.477 Million cell updates/sec

Title: US-09-837-138-1  
Perfect score: 4403  
Sequence: 1 ttcggcagcagcgcggttg.....accgcggtggagctccagct 4403

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sta:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4164.2	94.6	4406	9 AF058696	AF058696 Homo sapi
2	4123.4	93.6	4412	9 AF051334	AF051334 Homo sapi
3	4115.2	93.5	4388	6 CQ725600	CQ725600 Sequence
4	4108.4	93.3	4386	6 AR233599	AR233599 Sequence
5	4102	93.2	4466	9 BC071590	BC071590 Homo sapi
6	4038.4	91.7	4423	11 BV177869	BV177869 sqnm97272
7	1862.8	42.3	55134	9 AY566246	AY566246 Homo sapi
8	1862.8	42.3	56500	9 AB013139	AB013139 Homo sapi
9	1862.8	42.3	184919	9 AF069291	AF069291 Homo sapi
10	1862.8	42.3	331864	9 AF049895	AF049895 Homo sapi
11	1861.2	42.3	107549	9 AC074178	AC074178 Homo sapi
12	1836.4	41.7	2044	6 BD155805	BD155805 Primer fo
13	1836.4	41.7	2044	6 AX875862	AX875862 Sequence
14	1836.4	41.7	2044	9 AK001017	AK001017 Homo sapi
15	1416.4	32.2	2605	10 AF218575	AF218575 Rattus no
16	1405	31.9	2503	10 BC055061	BC055061 Mus muscu
17	1400.4	31.8	2484	10 AF076887	AF076887 Mus muscu
18	1398.8	31.8	2473	10 AF092840	AF092840 Mus muscu
19	1398.6	31.8	2542	10 BC044773	BC044773 Mus muscu

20	1395.6	31.7	2412	10 AB016988	AB016988 Mus muscu
21	1391	31.6	2518	10 BC003719	BC003719 Mus muscu
22	694	15.8	72910	2 AC023998	AC023998 Homo sapi
23	684	15.5	2533	5 AF230342	AF230342 Gallus ga
24	671	15.2	752	6 BD145652	BD145652 Primer fo
25	671	15.2	752	6 AX865590	AX865590 Sequence
26	564.6	12.8	622	6 BD179056	BD179056 Cancer as
27	552.2	12.5	646	6 BD265337	BD265337 Compounds
28	552.2	12.5	646	6 AR401323	AR401323 Sequence
29	552.2	12.5	646	6 AX192776	AX192776 Sequence
30	538.6	12.2	544	6 AX341490	AX341490 Sequence
31	520.4	11.8	543	6 BD151180	BD151180 Primer fo
32	520.4	11.8	543	6 AX871118	AX871118 Sequence
33	478.4	10.9	481	6 CQ681271	CQ681271 Sequence
34	456	10.4	2444	5 AY207367	AY207367 Xenopus l
35	455.6	10.3	2292	5 AV312176	AV312176 Xenopus l
36	438.2	10.0	439	6 CQ662253	CQ662253 Sequence
37	422.4	9.6	614	11 BV174416	BV174416 sqnm71594
38	358.4	8.1	396	6 AX067375	AX067375 Sequence
39	354.2	8.0	1811	5 CR848449	CR848449 Xenopus t
40	349.6	7.9	1666	5 BC082430	BC082430 Xenopus l
41	275	6.2	72910	2 AC023998	AC023998 Homo sapi
42	273.4	6.2	434	11 G66922	G66922 csnphb1-pc
43	264.4	6.0	421	11 G30107	G30107 human STS S
44	212.6	4.8	475	6 AX401127	AX401127 Sequence
45	210.8	4.8	238844	2 AC108261	AC108261 Rattus no

ALIGNMENTS

RESULT 1  
AF058696  
LOCUS AF058696 4406 bp mRNA linear PRI 22-MAY-2003  
DEFINITION Homo sapiens cell cycle regulatory protein p95 (NBS1) mRNA,  
Complete cds.  
ACCESSION AF058696  
VERSION AF058696.2 GI:30995499  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 4406)  
AUTHORS Carney, J.P., Maser, R.S., Olivares, H., Davis, E.M., Le Beau, M.,  
Yates, J.R., III, Hays, L., Morgan, W.F. and Petrini, J.H.  
TITLE The hMre11/hRad50 protein complex and Nijmegen breakage syndrome:  
linkage of double-strand break repair to the cellular DNA damage  
response  
JOURNAL Cell 93 (3), 477-486 (1998)  
MEDLINE 98250063  
PUBMED 9590181  
REFERENCE 2 (bases 1 to 4406)  
AUTHORS Carney, J.P., Maser, R.S., Olivares, H.A., Davis, E.M., Le Beau, M.,  
Yates, J.R., III, Hays, L., Morgan, W.F. and Petrini, J.H.J.  
TITLE Direct Submission  
JOURNAL Submitted (08-APR-1998) Genetics, University of Wisconsin, 445  
Henry Mall, Madison, WI 53706, USA  
COMMENT On May 22, 2003 this sequence version replaced gi:3098674.  
Sequence updated by GenBank staff.  
FEATURES  
Location/Qualifiers  
source  
1..4406  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8q21.3"  
1..4406  
/gene="NBS1"  
53..2317  
/gene="NBS1"  
/function="DNA repair"  
/note="nibrin"



```
/codon_start=1
/product="cell cycle regulatory protein p95"
/protein_id="AAC39752.1"
/db_xref="GI:3098675"
/translation="MKKLIPAAAGPAGGEPYRLLTGYEYVVRKNCALILINDQSI SRN
HAYLTAFNSVTNLSDTDEIPVLTKDLSKYGTFTVNEKMQNGFSRLTKSGDGIITFGVP
GSKFRIEYBLVACSCLDVSGKTAALNOAILQGGFTVNNWTECHLVMSVKVITIK
TICALICGRPIVKPVEYETBFLKAVQSKOPPOLESFVPLDPSIGSKNDVLSGRQER
KQIFKGTPIFLNAKHKKLSSAVFVGGEARLITEENSEEHNFFLAPGTCVVDGII
NSQTLIPDCQKWIQSIDMLQROGLRPIPEAEIGLAVIFMTTKNTCDQGHPSGLK
TTTTGSLSGQVSDSKJMPSPAFVNTTYYVADTESQADTWLDSRPKEIKVSKMEQK
FRMSQADATKVESCTSGNNNSMNTLAKMRI PNYSLSPTKLPISINKSKDRASQQ
QWISIRNYFOPSTKKEBERDENQMSCKSARIETSCSLEOTOPATPSLWKNKEOHL
SENEPVDTHSDNNLFTDOLKSIKNSASKSHAAEKLRSNKKREMDVIAIEDEVLQ
FKDTKELEIDVVKQBEDVNVNRKPRMDIETNDFSDPAVPESKISQENEIGKKR
ELKEDLSNAKEISQNDKLQDSEMLPKLLLTPEFRLSVIKNSTSNPSGINDDYQQL
KNPKFKKVTYPGAKLPHIIGGSLIAHHARKNTLBEWLRQEMEYVQNHAKESLSA
DDLFRYNPYLKRKR"
```

ORIGIN

Query Match		94.6%;	Score 4164.2;	DB 9;	Length 4406;
Best Local Similarity		97.6%;	Pred. No. 0;		
Matches 4303;		Conservative	7;	Mismatches	10;
				Indels	87;
				Gaps	5;
Qy	1	TTCCGCA	CAGAGCGCGGTTG	CACGT	CGGCCCCAGCCCTGAGAGCGCGGACCGGATGGAA 60
Db	1	TTCCGCA	CAGAGCGCGGTTG	CACGT	CGGCCCCAGCCCTGAGAGCGCGGACCGGATGGAA 60
Qy	61	ACTGCTG	CGCCCGCGCGCGCGG	CCCGCAGGAGGAGAACCATACAGACTTTTGA	CTGGCGTTGA 120
Db	61	ACTGCTG	CGCCCGCGCGCGCGG	CCCGCAGGAGGAGAACCATACAGACTTTTGA	CTGGCGTTGA 120
Qy	121	GTACGTT	GTGGAGGAAAAAC	TGTGCCATCTTAATTTGAAAAATGATCAGTCGATGACCGG	180
Db	121	GTACGTT	GTGGAGGAAAAAC	TGTGCCATCTTAATTTGAAAAATGATCAGTCGATGACCGG	180
Qy	181	AAATCAT	GTGCTGTGTTAACTGCTAACTTTCTGTAACCAACTCAGTCAACACAGATGAAAT	240	
Db	181	AAATCAT	GTGCTGTGTTAACTGCTAACTTTCTGTAACCAACTCAGTCAACACAGATGAAAT	240	
Qy	241	CCCTGTA	TATTAACATTAAGATAAATCTAAGTATGGTACCTTTGTTAAATGAGGAAAAAAT	300	
Db	241	CCCTGTA	TATTAACATTAAGATAAATCTAAGTATGGTACCTTTGTTAAATGAGGAAAAAAT	300	
Qy	301	GCAGAT	TGGCTTTCCGGAACCTTTGAAGTCGGGGGATGGTATTACTTTTGGAGTGTTCG	360	
Db	301	GCAGAT	TGGCTTTCCGGAACCTTTGAAGTCGGGGGATGGTATTACTTTTGGAGTGTTCG	360	
Qy	361	AAGTAAAT	TGAGATAGAGTATGAGCCTTTGGTTGTCATGCTCTCTTGTGTTAGATGCTC	420	
Db	361	AAGTAAAT	TGAGATAGAGTATGAGCCTTTGGTTGTCATGCTCTCTTGTGTTAGATGCTC	420	
Qy	421	TGGGAAAT	CTGCTTTTAAATCAAGCTATAATTGCAACTTGGAGGATTTACTGTAAACAAATG	480	
Db	421	TGGGAAAT	CTGCTTTTAAATCAAGCTATAATTGCAACTTGGAGGATTTACTGTAAACAAATG	480	
Qy	481	GACAGAA	TGACATGACATCTGTCATGGTATCAGTGAAGTACCATTAAACATATG	540	
Db	481	GACAGAA	TGACATGACATCTGTCATGGTATCAGTGAAGTACCATTAAACATATG	540	
Qy	541	TGCAC	TCTTTTGGAGCTCCAATGTAAAGCCAGAAATATTTTACTGAAATTCCTGAAAGC	600	
Db	541	TGCAC	TCTTTTGGAGCTCCAATGTAAAGCCAGAAATATTTTACTGAAATTCCTGAAAGC	600	
Qy	601	AGTTCA	GTCAGTCCAGAGCAGCTCCACAAATTTGAAAGTTTTTACCACCTCTTGATGAACC	660	
Db	601	AGTTCA	GTCAGTCCAGAGCAGCTCCACAAATTTGAAAGTTTTTACCACCTCTTGATGAACC	660	
Qy	661	ATCTAT	TGGAGTAAATATGTTGATCTGTACAGCCGCGAGGAAAGAAACAAATCTTCAA	720	
Db	661	ATCTAT	TGGAGTAAATATGTTGATCTGTACAGCCGCGAGGAAAGAAACAAATCTTCAA	720	
Qy	721	AGGGAAA	CAATTTATATTTTTTGAATGCCAAACAGCATAGAAATTGAGTTCCGCGAGTTGT	780	

DB	721	AGGGAAA	CAATTTATATTTTTTGAATGCCAAACAGCATAGAAATTGAGTTCCGCGAGTTGT 780
QY	781	CTTTGGA	GGTGGGAGCTAGGTTGATTAACAGAGAGAAATGAAGAAACATAAATTTCTT 840
DB	781	CTTTGGA	GGTGGGAGCTAGGTTGATTAACAGAGAGAAATGAAGAAACATAAATTTCTT 840
QY	841	TTTTGCT	CCGGAAACGTGTGTTGTTGATACAGGAATACAAACTCACAGACCTTAATTC 900
DB	841	TTTTGCT	CCGGAAACGTGTGTTGTTGATACAGGAATACAAACTCACAGACCTTAATTC 900
QY	901	TGACTGT	CAGAAAGAAATGGAATTCAGTCAATATGATATGCTCCAAAGGCAAGGCTTAG 960
DB	901	TGACTGT	CAGAAAGAAATGGAATTCAGTCAATATGATATGCTCCAAAGGCAAGGCTTAG 960
QY	961	ACTAT	TCTTGAAGCAGAAATTTGGATTTGGCGGTGATTTTTCATGCTACAAAGAAATAC 1020
DB	961	ACTAT	TCTTGAAGCAGAAATTTGGATTTGGCGGTGATTTTTCATGCTACAAAGAAATAC 1020
QY	1021	TGATCCT	CAGGGCCATCCAGGTACAGGATTAAGACAACTCCAGGACCAAGCCTTTC 1080
DB	1021	TGATCCT	CAGGGCCATCCAGGTACAGGATTAAGACAACTCCAGGACCAAGCCTTTC 1080
QY	1081	ACAAG	CGTGTGATGAAAACTAATGCAAGCGCCCCAGTGAACTACAACTA 1140
DB	1081	ACAAG	CGTGTGATGAAAACTAATGCAAGCGCCCCAGTGAACTACAACTA 1140
QY	1141	CGTAG	CTGACACAGAAATCAGAGCAAGCAGATACATGGGATTTGAGTAAAGGCCAAAGA 1200
DB	1141	CGTAG	CTGACACAGAAATCAGAGCAAGCAGATACATGGGATTTGAGTAAAGGCCAAAGA 1200
QY	1201	AATCA	AGTCTCCAAAATGGAAACAAATTCAGAAATGCTTTCACAGAGCGCACCTG 1260
DB	1201	AATCA	AGTCTCCAAAATGGAAACAAATTCAGAAATGCTTTCACAGAGCGCACCTG 1260
QY	1261	AAAGG	AGTCTCCAAAACAAAGCTCTAATAATAGTATGTTGATCAAAATCTTTGGCTAA 1320
DB	1261	AAAGG	AGTCTCCAAAACAAAGCTCTAATAATAGTATGTTGATCAAAATCTTTGGCTAA 1320
QY	1321	GATGA	GAATCCCAAATCTATCAGCTTTCCAACTAAATTTGCCAAGTATAATAAAGTAA 1380
DB	1321	GATGA	GAATCCCAAATCTATCAGCTTTCCAACTAAATTTGCCAAGTATAATAAAGTAA 1380
QY	1381	AGAT	AGGCTCTCAGCAGCAGCAGCAACCTCCATCAGAAACTACTTTTCAGCCGCTAC 1440
DB	1381	AGAT	AGGCTCTCAGCAGCAGCAGCAACCTCCATCAGAAACTACTTTTCAGCCGCTAC 1440
QY	1441	CAAAA	AAAAAGGGAAGGATGAAGAAATCAAGAAATGTCTTCATGCAAAATCAGCAAGAT 1500
DB	1441	CAAAA	AAAAAGGGAAGGATGAAGAAATCAAGAAATGTCTTCATGCAAAATCAGCAAGAT 1500
QY	1501	AGAAA	CGTCTGTTCTCTTTTAGAACAAACACAACTGTCACCTCACTTGTGGAAAA 1560
DB	1501	AGAAA	CGTCTGTTCTCTTTTAGAACAAACACAACTGTCACCTCACTTGTGGAAAA 1560
QY	1561	TAAG	GAGCAGCATCTATCTGAGAAATGAGCTGTGACACAACTCAGACAAATACTATT 1620
DB	1561	TAAG	GAGCAGCATCTATCTGAGAAATGAGCTGTGAGCAAACTCAGACAAATACTATT 1620
QY	1621	TACAG	TACAGATTTTAAATCTATTGTGAAAAATCTGCCAGTAAATCTCATGCTGCAGA 1680
DB	1621	TACAG	TACAGATTTTAAATCTATTGTGAAAAATCTGCCAGTAAATCTCATGCTGCAGA 1680
QY	1681	AAAG	CTAAGTAAATTAAGGAAATGGATGATGGCCATAGAGATGAAGTATT 1740
DB	1681	AAAG	CTAAGTAAATTAAGGAAATGGATGATGGCCATAGAGATGAAGTATT 1740
QY	1741	GGAA	CAGTTATTCAAGGACACAAACACAGATTTAGAAATTTGATGTGAAAGTTCAAAA 1800
DB	1741	GGAA	CAGTTATTCAAGGACACAAACACAGATTTAGAAATTTGATGTGAAAGTTCAAAA 1800
QY	1801	GGAG	GAAGATGTCAAATGTTAGAAAAGGCCAAGGATGGATATAGAAAACAAATGACACT 1860



Db 1801 GGAGGAAGATGTCATGTTTGAAGAAAGGCCAAGGATGGAATATAGAAACAAATGACACTTT 1860  
Qy 1861 CAGTGATGAAGCAGTACAGAAAGTAGCAAAATATCTCAAGAAATGAAATTTGGGAAGAA 1920  
Db 1861 CAGTGATGAAGCAGTACAGAAAGTAGCAAAATATCTCAAGAAATGAAATTTGGGAAGAA 1920  
Qy 1921 ACGTGAACCTCAAGGAAGACTCCTATGAGTCACTATGAGTCACTAAAGAAATATCTAAATGACAAACT 1980  
Db 1921 ACGTGAACCTCAAGGAAGACTCCTATGAGTCACTAAAGAAATATCTAAATGACAAACT 1980  
Qy 1981 TCAGGATGATGAGTGAAGTCTCCAAAAAGCTGTTATTGACCTGAAATTTAGATCACTGGT 2040  
Db 1981 TCAGGATGATGAGTGAAGTCTCCAAAAAGCTGTTATTGACCTGAAATTTAGATCACTGGT 2040  
Qy 2041 GATTAAAAACTCTACTCTTCAGAAATCCGCTCGGCATAAAATGATGATTATGGTCAACTAAA 2100  
Db 2041 GATTAAAAACTCTACTCTTCAGAAATCCGCTCGGCATAAAATGATGATTATGGTCAACTAAA 2100  
Qy 2101 AAATTTCAAGAAATTCAAAAAGTCAATATCTCGAGGAGGAAACTTCCACACATCAT 2160  
Db 2101 AAATTTCAAGAAATTCAAAAAGTCAATATCTCGAGGAGGAAACTTCCACACATCAT 2160  
Qy 2161 TGAGGATCAGATCTAATAGCTCATCATGCTCGAAAGAAATACAGAACTAAGAGTGGCT 2220  
Db 2161 TGAGGATCAGATCTAATAGCTCATCATGCTCGAAAGAAATACAGAACTAAGAGTGGCT 2220  
Qy 2221 AAGCAGGAAATGGAGTACAAAATCAACATGCAAAAGAGAGTCTCTTCTGATGATCT 2280  
Db 2221 AAGCAGGAAATGGAGTACAAAATCAACATGCAAAAGAGAGTCTCTTCTGATGATCT 2280  
Qy 2281 TTTTAGATCAATCTCTTATTTAAAAAGGAGGAAGATAACTGAGGATTTAAAAAGAGCCA 2340  
Db 2281 TTTTAGATCAATCTCTTATTTAAAAAGGAGGAAGATAACTGAGGATTTAAAAAGAGCCA 2340  
Qy 2341 TGAAAACTTCTCTAGTACATCTCTCAGGCCAACAGGTTATATGAATATATAGTG 2400  
Db 2341 TGAAAACTTCTCTAGTACATCTCTCAGGCCAACAGGTTATATGAATATATAGTG 2400  
Qy 2401 TATAGAAGCGATTTAAGTTACAAATGTTTATGCGCTAAATTTATTAATAAAATGCACAA 2460  
Db 2401 TATAGAAGCGATTTAAGTTACAAATGTTTATGCGCTAAATTTATTAATAAAATGCACAA 2460  
Qy 2461 AACTTTGATCTTTTGTATGATGAACAAATGTTTGTCTGTTTCAAGGCTTTGTCAATGCAAT 2520  
Db 2461 AACTTTGATCTTTTGTATGATGAACAAATGTTTGTCTGTTTCAAGGCTTTGTCAATGCAAT 2520  
Qy 2521 CTTTTTTTCAATTTTAAATGCTTTTGTGTTTATTAATAGTATATAGTACAGTTCCAA 2580  
Db 2521 CTTTTTTTCAATTTTAAATGCTTTTGTGTTTATTAATAGTATATAGTACAGTTCCAA 2580  
Qy 2581 ATTCTAAATRTAGCTAAGGTAAAGGACTAAAGTCACCTTCCACCAATTTGCTTAGCTACT 2640  
Db 2581 ATTCTAAATRTAGCTAAGGTAAAGTAAAGTCAACCTTCCACCAATTTGCTTAGCTACT 2640  
Qy 2641 ----- 2640  
Db 2640 TGCTTCCCTTCAGAAAAAATTCATGATCTCATTTCTTATGAATCTTTCCAGGGATTTT 2699  
Qy 2641 -----TATTTTAAATAATTTCTCTACACAAATGATAGCATACATA 2681  
Db 2700 TGAGTCCCTATTCAAATTTCTTATTTTAAATAATTTCTCTACACAAATGATAGCATACATA 2759  
Qy 2682 TGCAAGTGTCTACACCTTGCTTTTCTTACTTAAAGATTAAGAAATTTATAGGAATATCAAT 2741  
Db 2760 TGCAAGTGTCTACACCTTGCTTTTCTTACTTTAGT-AGANTAAATTTATAGGAATATCAAT 2818  
Qy 2742 ATAATGTTTTTAATATTTTTTTCTTTTCCATTTAGTGTAGTCTTACCTAACTCTGGTGA 2801  
Db 2819 ATAATGTTTTTAATATTTTTTTCTTTTCCATTTAGTGTAGTCTTACCTAACTCTGGTGA 2878  
Qy 2802 TCCAAACAAATGGCTTCAGTGTGCGATGTGCACCTACATGTTTCTAGTACTAGAAA 2861  
Db 2879 TCCAAACAAATGGCTTCAGTGTGCGATGTGCACCTACATGTTTCTAGTACTAGAAA 2938

Qy 2862 CTGAAGACCATGTGGAGACTTTCATCAAAATCATGGGTTTAGTTTTTACCAGAAATGGAAGAC 2921  
Db 2939 CTGAAGACCATGTGGAGACTTTCATCAAAATCATGGGTTTAGTTTTTACCAGAAATGGAAGAC 2998  
Qy 2922 CTGTACCCCTTTTGGTGTCTTACTAGCTGGTGGTGTCTGTGTTTGGAGCTTATTAG 2981  
Db 2999 CTGTACCCCTTTTGGTGTCTTACTAGCTGGTGGTGTCTGTGTTTGGAGCTTATTAG 3058  
Qy 2982 AGTCCTAGTGTCTTACTTATTAAGTAGAAATGGTGGAGATTTGTTTCTTTTCTACCKTA 3041  
Db 3059 AGTCCTAGTGTCTTACTTATTAAGTAGAAATGGTGGAGATTTGTTTCTTTTCTACCTTA 3118  
Qy 3042 AAGGGAGATGGTAAGAAAACAATGAATGTCTTTTTCATAAATTTTATTTGACAAGTGAATTC 3101  
Db 3119 AAGGGAGATGGTAAGAAAACAATGAATGTCTTTTTCATAAATTTTATTTGACAAGTGAATTC 3178  
Qy 3102 AGTCTGTCTTCAAAATATATTTTATGATCTGATCCAGCAAGAGGGAGTTCAGTTC 3161  
Db 3179 AAGTCTGTCTTCAAAATATATTTTATGATCTGATCCAGCAAGAGGGAGTTCAGTTC 3238  
Qy 3162 AAGAGTCACTACAACTGATTTAGTTTGTAGAGAAATGAGAAATGGAACAGTGAGGAATGGA 3221  
Db 3239 AAGAGTCACTACAACTGATTTAGTTTGTAGAGAAATGAGAAATGGAACAGTGAGGAATGGA 3298  
Qy 3222 GGCCATATTTTCCATGACTTCCCTTTGTAACAGAGCAACAGAGGCAAGAGGCTGGCC 3281  
Db 3299 GGCCATATTTTCCATGACTTCCCTTTGTAACAGAGCAACAGAGGCAAGAGGCTGGCC 3358  
Qy 3282 TCTACATCACTCACCCTTCCAAATCTTGTGGAAGTGCATCTACTTTCAGCAACCAATTT 3341  
Db 3359 TCTACATCACTCACCCTTCCAAATCTTGTGGAAGTGCATCTACTTTCAGCAACCAATTT 3418  
Qy 3342 AACTTACTTCCAAAGTCTGCTGCTGAGGTGGAATCCAGCTGCAAGGGAGTTAGGGA 3401  
Db 3419 AACTTACTTCCAAAGTCTGCTGCTGAGGTGGAATCCAGCTGCAAGGGAGTTAGGGA 3478  
Qy 3402 AATGAAGGTCTTTTTTAAAAAGCTTCTCAGCCCTTCTAGGCAACAGAAATGGGTGAGCC 3461  
Db 3479 AATGAAGGTCTTTTTTAAAAAGCTTCTCAGCCCTTCTAGGCAACAGAAATGGGTGAGCC 3538  
Qy 3462 AATTCGAATTTCTACTACAGGATTTAGACCAAGTTAGATTTGAAATATTTATAGAGAG 3521  
Db 3539 AATTCGAATTTCTACTACAGGATTTAGACCAAGTTAGATTTGAAATATTTATAGAGAG 3598  
Qy 3522 TTATGAACACTTAAATTTATGATAGTGTATGATGAAATGGATAGAAATGGGTACTTTAGA 3581  
Db 3599 TTATGAACACTTAAATTTATGATAGTGTATGATGAAATGGATAGAAATGGGTACTTTAGA 3658  
Qy 3582 AGTAGAATTTGACAGGCTATTTAGTTGATGAAATGGAGTCAATTTGAGTCTTATTAATAGCC 3641  
Db 3659 AGTAGAATTTGACAGGCTATTTAGTTGATGAAATGGAGTCAATTTGAGTCTTATTAATAGCC 3718  
Qy 3642 ATGTATCAATAATTAACCAAGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGG 3701  
Db 3719 ATGTATCAATAATTAACCAAGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGG 3778  
Qy 3702 AATATAATGACAGATTAATAATTTGTTCTGTCATGCCCAATCCCTTTCTTAAGGAAGA 3761  
Db 3779 AATATAATGACAGATTAATAATTTGTTCTGTCATGCCCAATCCCTTTCTTAAGGAAGA 3838  
Qy 3762 CTGCCCCCTACTATAGCAGTTTTTATATTTGTCATTTTATGAATATATGAATGAGGAGTTTC 3821  
Db 3839 CTGCCCCCTACTATAGCAGTTTTTATATTTGTCATTTTATGAATATATGAATGAGGAGTTTC 3898  
Qy 3822 TGGTACCTCTGTCTTTTCAAAATATTTGGGTGTGTGTCAGTATTTTCCCTTTTAAACCMT 3881  
Db 3899 TGGTACCTCTGTCTTTTCAAAATATTTGGGTGTGTGTCAGTATTTTCCCTTTTAAACCAT 3958  
Qy 3882 TCCCAATTCGGGTGTAGTGGATGTTTCCATTTTGGGTGTGTTTAAATTTGATATATCCCTGAT 3941  
Db 3959 T-CCAATCGGTGTAGTGTATG- - - - -TTTCATTTTGGTTTAAATTTGATATATCCCTGAT 4012



QY 3942 AGCTATAATTGGGTCATAGAAATCTTTATACATTTCTAGATCAAGTCTCTTGCGGATA 4001  
|||||  
Db 4013 AGCTATAATTGGGTCATAGAAATCTTTATACATTTCTAGATCAAGTCTCTTGCGGATA 4072  
|||||  
QY 4002 TACGTATTGAGATATTACACCTAGTCTGTGGCTTGACTGTTCTTTTATGTCTTTTGTATG 4061  
|||||  
Db 4073 TATGTTATGAGATATTACACCTAGTCTGTGGCTTGACTGTTCTTTTATGTCTTTTGTATG 4132  
|||||  
QY 4062 AATAGAAGTTTTAAATTTTGACAGGTCAAAATTTATTTTCTTTTGTGATATTTT 4121  
|||||  
Db 4133 AATAGAAGTTTTAAATTTTGACAGGTCAAAATTTATTTTCTTTTGTGATATTTT 4192  
|||||  
QY 4122 TCTCTCAATTTAAACCCCAAGATTTTCAGATATTCTGCTATTATATATAAACTTTATATTT 4181  
|||||  
Db 4193 TCTCTCAATTTAAACCCCAAGATTTTCAGATATTCTGCTATTATATATAAACTTTATATTT 4252  
|||||  
QY 4182 TTATATTGTGATCTACCTTTGAATGATATGATGTTGTGAATATGATCGAGGTTCTT 4241  
|||||  
Db 4253 TTATATTGTGATCTACCTTTGAATGATATGATGTTGTGAATATGATCGAGGTTCTT 4312  
|||||  
QY 4242 TTTTCCGCCATACAGATATCCAGTCATTTGTAACACTGTTTATTTGAAGAAATATCCTTT 4301  
|||||  
Db 4313 TTTTCCGCCATACAGATATCCAGTCATTTGTAACACTGTTTATTTGAAGAAATATCCTTT 4372  
|||||  
QY 4302 CCTCATTAATTAACCTTTGCCAATTAGT 4328  
|||||  
Db 4373 CCTCATTAATTAACCTTTGCCAATTAGT 4399  
|||||

RESULT 2  
AF051334 4412 bp mRNA linear PRI 09-MAY-1998  
LOCUS Homo sapiens nibrin (NBS) mRNA, complete cds.  
DEFINITION AF051334  
ACCESSION AF051334.1 GI:3126794  
VERSION  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Eukaryota; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 4412)  
AUTHORS Varon,R., Vissinga,C., Platzer,M., Cerosaletti,K.M.,  
Chrzanowska,K.H., Saar,K., Beckmann,G., Seemanova,E., Cooper,P.R.,  
Nowak,N.J., Stumm,M., Weemaes,C.M.R., Gatti,R.A., Wilson,R.K.,  
Digweed,M., Rosenchal,A., Sperling,K., Concannon,P. and Reis,A.  
Nibrin, a novel DNA double-strand break repair protein, is mutated  
in Nijmegen breakage syndrome  
Cell 93 (3), 467-476 (1998)  
JOURNAL 98250062  
MEDLINE 9590180  
PUBMED  
REFERENCE 2 (bases 1 to 4412)  
AUTHORS Varon,R. and Platzer,M.  
Direct Submission  
TITLE Submitted (26-FEB-1998) Genome Analysis, Institute of Molecular  
Biotechnology, Beutenbergstr.11, Jena 07745, Germany  
FEATURES  
source  
1. .4412  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
gene  
1. \_4412  
/gene="NBS"  
CDS  
52..2316  
/gene="NBS"  
/notes="mutated in Nijmegen Breakage syndrome"  
/codon\_start=1  
/evidence=not experimental  
/product="nibrin"  
/protein\_id="AAC39732.1"  
/db\_xref="GI:3126795"  
translation="MMKLLPAPAGPAGGEPYRLLTGVYVYVGRKNCAILIENDQISRN  
HVLITANFVNLSQSDIEIPVLTLDKNSKYIGTFVNEKMQNGFSRLTKSGDGTTFGVF  
GSKFRIEYPLVACSSCLDVSQKLTALNQAILQGGFTVNNTEECTHLVWVSVKYTIK

TICALICGRPIVKPYFTFLKAVBSKKOPPOIESFYPPFLDEPSIGSKNVDLSGRQER  
KOIPKGTFFILNAKQHKLASAVVGGGEARLITEENEENHFFLAEGTCVVDGTGIT  
NSQTLIPQCKKNVDSIMDLQROQLRPIPEAIEGLAVIPMTKNYCDPQHPSTGLK  
TTTPPSUSQGVSDKELMPVNVNTTIVADTESEQADTDLSERPKEIKVSKWEQK  
FRLMSQDAPTQKESCTSSNNNSMVSNTLAKMRIPNYQLSPTKLPSINKSKDRASQQQ  
QTNSTRNTFQPSKTRDERDENQESSCKSARIEISCSLLEQTPATPSLWKNKEQHL  
SENEPVDNTSNLFTDLDKSI VKNASAKSHAARKSNKKREMDVIAIEDEVLEQL  
PKDTPELEIDVKQKOBEDVNVKRPRMDIETNDTPFSDKAEVPSKISOEINEICKRR  
ELKEDLSAKSEIENNDKLODDSEMLPKLLLTPEPSLVIKNSTSRNPSGINDDYGOL  
KNFKPKVTTYPGAGKPHIIGGSLIIAHARKNTELEBWLKQEMEVONQHAKEESLA  
DDLFRYNPYLTKRRR"  
polyA\_signal  
2446..2451  
/gene="NBS"  
/evidence=experimental  
polyA\_signal  
4392..4397  
/gene="NBS"  
/evidence=experimental  
ORIGIN  
Query Match 93.6%; Score 4123.4; DB 9; Length 4412;  
Best Local Similarity 97.1%; Pred. No. 0;  
Matches 4287; Conservative 7; Mismatches 28; Indels 91; Gaps 6;  
QY 13 CGCGGTTGCACGTCGCGCCCGCCAGCCCTGAGGAGCCGACCGATGTGGAAACTCTGTCGCCG 72  
|||||  
Db 12 CGCGGTTGCACGTCGCGCCCGCCAGCCCTGAGGAGCCGACCGATGTGGAAACTCTGTCGCCG 71  
|||||  
QY 73 CGCGGCGCGGCGAGGAGAAACCATACAGACTTTTGTGACTGCGGTTGAGTACGTTGTGG 132  
|||||  
Db 72 CGCGGCGCGGCGAGGAGAAACCATACAGACTTTTGTGACTGCGGTTGAGTACGTTGTGG 131  
|||||  
QY 133 AAGGAAAACTCTGCGCATCTTAATTTGAAATCATCATCGATCGATCGACGCGAAATCATGCTGT 192  
|||||  
Db 132 AAGGAAAACTCTGCGCATCTTAATTTGAAATCATCATCGATCGATCGACGCGAAATCATGCTGT 191  
|||||  
QY 193 GTTAACTGCTAACTTTTCTGTAAACCAACCTGAGTCAACACAGATGAAATCCCTGTATTGAC 252  
|||||  
Db 192 GTTAACTGCTAACTTTTCTGTAAACCAACCTGAGTCAACACAGATGAAATCCCTGTATTGAC 251  
|||||  
QY 253 ATTTAAAGATAATTTCTAAGTATGATGTCCTTTGTTAATGAGGAAAAATGACAGAAATGGCTT 312  
|||||  
Db 252 ATTTAAAGATAATTTCTAAGTATGATGTCCTTTGTTAATGAGGAAAAATGACAGAAATGGCTT 311  
|||||  
QY 313 TTCCCGAACTTTGAAGTCGGGGATGATTAATTTTTCGAGTGTGTTGGAAGTAAATTCAG 372  
|||||  
Db 312 TTCCCGAACTTTGAAGTCGGGGATGATTAATTTTTCGAGTGTGTTGGAAGTAAATTCAG 371  
|||||  
QY 373 AATAGAGATGAGCCCTTTGGTTGCAATGCTCTTTCTTTTGTAGATGCTCTGCGGAAAACTGC 432  
|||||  
Db 372 AATAGAGATGAGCCCTTTGGTTGCAATGCTCTTTCTTTTGTAGATGCTCTGCGGAAAACTGC 431  
|||||  
QY 433 TTTAAATCAAGCTATATTGCAACTTTGGAGATTTACTGTAAACAATTTGACAGAGAATG 492  
|||||  
Db 432 TTTAAATCAAGCTATATTGCAACTTTGGAGATTTACTGTGTAAACAATTTGACAGAGAATG 491  
|||||  
QY 493 CACTCACCTTGTTCATGTTATCAGTCAAAAGTTACCAATTAACCAATATGTGCACTCATTTG 552  
|||||  
Db 492 CACTCACCTTGTTCATGTTATCAGTCAAAAGTTACCAATTAACCAATATGTGCACTCATTTG 551  
|||||  
QY 553 TGGAGCTCCAATTTGTAAGCCAGAAATATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAA 612  
|||||  
Db 552 TGGAGCTCCAATTTGTAAGCCAGAAATATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAA 611  
|||||  
QY 613 GAAGCAGCTCCACAAATTTGAAGTTTATACCACCTCTTGTATGAACCATCTATTGGAG 672  
|||||  
Db 612 GAAGCAGCTCCACAAATTTGAAGTTTATACCACCTCTTGTATGAACCATCTATTGGAG 671  
|||||  
QY 673 TAAAAATGTTGATCTGTCAAGCAGCGAGGAGAAAAAATAATCTTCAAGGAAAAACATT 732  
|||||  
Db 672 TAAAAATGTTGATCTGTCAAGCAGCGAGGAGAAAAAATAATCTTCAAGGAAAAACATT 731  
|||||  
QY 733 TATATTTTGAATGCCAAACACGATAAGAAATTTAGTTCCGAGTTGCTTTTGGAGGTGG 792  
|||||



Db	732	TATATTTTTGAAATGCCAAACAGCATAAAGAAATGAGTTCCGCGAGTTGCTTTTGGAGGTGG	791
Qy	793	GGAAAGCTAGGTTTGATAACAGAAAGAGAAATCAAGAAACAATAATTTCTTTTGGCTCCGGG	852
Db	792	GGAAAGCTAGGTTTGATAACAGAAAGAGAAATCAAGAAACAATAATTTCTTTTGGCTCCGGG	851
Qy	853	AACGTGTGTTGTGATACAGGAATAACAAACTCACAGACCTTTAAATTCCTGACTGTGAGAA	912
Db	852	AACGTGTGTTGTGATACAGGAATAACAAACTCACAGACCTTTAAATTCCTGACTGTGAGAA	911
Qy	913	GAAATGGATTACGTCAATTAATGSGATATGCTCCAAAGGCAAGGTCTTAGACCTATTCTCTGA	972
Db	912	GAAATGGATTACGTCAATTAATGSGATATGCTCCAAAGGCAAGGTCTTAGACCTATTCTCTGA	971
Qy	973	AGCAGAAATTTGATTTGGCGGTGATTTTTCATGACTACAAAGAAATTACTGTGATCTCTCAGGG	1032
Db	972	AGCAGAAATTTGATTTGGCGGTGATTTTTCATGACTACAAAGAAATTACTGTGATCTCTCAGGG	1031
Qy	1033	CCATCCCATGACAGATTAAAGACAACAACTCCAGGACCAAGCCCTTTCAAGGCGGTGTC	1092
Db	1032	CCATCCCATGACAGATTAAAGACAACAACTCCAGGACCAAGCCCTTTCAAGGCGGTGTC	1091
Qy	1093	AGTTGATGAAAACCTAAATGCCAAGCGCCCGAGTGAACTACAACTACGTAGCTGACAC	1152
Db	1092	AGTTGATGAAAACCTAAATGCCAAGCGCCCGAGTGAACTACAACTACGTAGCTGACAC	1151
Qy	1153	AGAAATCAGACGACAGATACATGCGGATTTGAGTGAAAGGCCAAAGAAATCAAAAGTCTC	1212
Db	1152	AGAAATCAGACGACAGATACATGCGGATTTGAGTGAAAGGCCAAAGAAATCAAAAGTCTC	1211
Qy	1213	CAAAATGGAACAAAAAATTCAGAAATCTTTCAACAAGCCACCTGTAAGAGGAGTCTTG	1272
Db	1212	CAAAATGGAACAAAAAATTCAGAAATCTTTCAACAAGCCACCTGTAAGAGGAGTCTTG	1271
Qy	1273	CAAAACAAGCTCTAATAATATGATGCGTATCAAACTATTGGCTAAGATGAGAAATCCC	1332
Db	1272	CAAAACAAGCTCTAATAATATGATGCGTATCAAACTATTGGCTAAGATGAGAAATCCC	1331
Qy	1333	AAACTATCAGCTTTTCCACCAACTAAATTCGCAAGTATAAATAAAAGTAAAGATAGCGTTC	1392
Db	1332	AAACTATCAGCTTTTCCACCAACTAAATTCGCAAGTATAAATAAAAGTAAAGATAGCGTTC	1391
Qy	1393	TCAGCAGCAGCAGACCAACTCCATCAGAAACTACTTTTCAGCCGCTCTACCAAAAAAGGGA	1452
Db	1392	TCAGCAGCAGCAGACCAACTCCATCAGAAACTACTTTTCAGCCGCTCTACCAAAAAAGGGA	1451
Qy	1453	AAGGATGAGAAATCAAGAAATGCTTCTATGCAATCAGCAGAGAAATAGAAACGTCTTG	1512
Db	1452	AAGGATGAGAAATCAAGAAATGCTTCTATGCAATCAGCAGAGAAATAGAAACGTCTTG	1511
Qy	1513	TTCTCTTTTAGAAACAAACACAACTGCTACACCTCTATTGGAATAATAGGAGCAGCA	1572
Db	1512	TTCTCTTTTAGAAACAAACACAACTGCTACACCTCTATTGGAATAATAGGAGCAGCA	1571
Qy	1573	TCTATCTGAGATGAGCCTGTGGAACAATACTCAGACAATAACTTATTTACAGATACAGA	1632
Db	1572	TCTATCTGAGATGAGCCTGTGGAACAATACTCAGACAATAACTTATTTACAGATACAGA	1631
Qy	1633	TTTAAATCTATTGTGAAAAATTTCTGCCAGTAAATCTCATGCTGCAAGAAAGCTAAGATC	1692
Db	1632	TTTAAATCTATTGTGAAAAATTTCTGCCAGTAAATCTCATGCTGCAAGAAAGCTAAGATC	1691
Qy	1693	AAATAAAAAAGGGAATGGAATGATGCGCCATAGAGATCAAGTATTGGAAACGATTATT	1752
Db	1692	AAATAAAAAAGGGAATGGAATGATGCGCCATAGAGATCAAGTATTGGAAACGATTATT	1751
Qy	1753	CAAGGACACAAAACAGAGTTAGAAATTTGATGTAAGCTTCAAAAACAGGAGGAAGTGT	1812
Db	1752	CAAGGACACAAAACAGAGTTAGAAATTTGATGTAAGCTTCAAAAACAGGAGGAAGTGT	1811
Qy	1813	CAATCTTTAGAAAAGGCCCAAGATGGAATATAGAAACAAATGACATTTCTAGTGTAGAAC	1872
Db	1812	CAATCTTTAGAAAAGGCCCAAGATGGAATATAGAAACAAATGACATTTCTAGTGTAGAAC	1871

1873	QY	AGTACCGAAGAGTAGCAAAATATCTCAAGAAAATGAAATTTGGGAGAGAAACGTGAACCTCAA	1931
1872	DB	AGTACCGAAGAGTAGCAAAATATCTCAAGAAAATGAAATTTGGGAGAGAAACGTGAACCTCAA	1931
1933	QY	GGAAAGACTCACATGGCTCAGCTAAAGAAAATATCTTAACAAATGACAAACCTTCAGGATGATAG	1992
1932	DB	GGAAAGACTCACATGGCTCAGCTAAAGAAAATATCTTAACAAATGACAAACCTTCAGGATGATAG	1991
1993	QY	TGAGATGCTTCCAAAAAGCTGTATTTGACTGAATTTTAGATCACTGGTGATTTAAAACTC	2052
1992	DB	TGAGATGCTTCCAAAAAGCTGTATTTGACTGAATTTTAGATCACTGGTGATTTAAAACTC	2051
2053	QY	TACTTTCAGAAATCCGCTCGCATAAATGATGATTATGGTCAACTAAAAAATTTCAAGAA	2112
2052	DB	TACTTTCAGAAATCCATCTCGCATAAATGATGATTATGGTCAACTAAAAAATTTCAAGAA	2111
2113	QY	ATTCAAAAAGGTCAATATCTCTGGAGCAGGAAAACTTTCCAACATCATTTGGAGGATCAGA	2172
2112	DB	ATTCAAAAAGGTCAATATCTCTGGAGCAGGAAAACTTTCCAACATCATTTGGAGGATCAGA	2171
2173	QY	TCTAATAGCTCATCATGCTCGAAGAAATACAGAACTAGAGAGTGGCTAAGGCAGGAAAT	2232
2172	DB	TCTAATAGCTCATCATGCTCGAAGAAATACAGAACTAGAGAGTGGCTAAGGCAGGAAAT	2231
2233	QY	GGAGGTACAAAATCAAATGCAAAAGAAAGAGTCTCTTGCTGATCATCTTTTAGATACAA	2292
2232	DB	GGAGGTACAAAATCAAATGCAAAAGAAAGAGTCTCTTGCTGATCATCTTTTAGATACAA	2291
2293	QY	TCCTTATTTAAAAAGGAGAGATAACTGAGGAATTTTAAAAAGAACCCATCGAAAACTTC	2352
2292	DB	TCCTTATTTAAAAAGGAGAGATAACTGAGGAGTTTTAAAAAGAACCCATCGAAAACTTC	2351
2353	QY	CTAGTAAGCATCTACTTTTCAGCCCAACAAGGTTATATGAATATATAGTGTATAGAAGCAT	2412
2352	DB	CTAGTAAGCATCTACTTTTCAGCCCAACAAGGTTATATGAATATATAGTGTATAGAAGCAT	2411
2413	QY	TTAAGTTACAATGTTTTATGGCCTPAAATTTTAAATAAAAATGCACAAAACCTTTGATTC	2472
2412	DB	TTAAGTTACAATGTTTTATGGCCTPAAATTTTAAATAAAAATGCACAAAACCTTTGATTC	2471
2473	QY	TTTGTATGTAAACATTCGTTGTCTGTTTTTCAGCCTTTGTTCATTCGATCTTTTTCAT	2532
2472	DB	TTTGTATGTAAACATTCGTTGTCTGTTTTTCAGCCTTTGTTCATTCGATCTTTTTCAT	2531
2533	QY	TTTAAATGCTGTTTTGTTTTTAAATAGTAAATATAGTCACAGTTCAAAATCTCAAAATPTA	2592
2532	DB	TTTAAATGCTGTTTTGTTTTTAAATAGTAAATATAGTCACAGTTCAAAATCTCAAAATPTA	2591
2593	QY	CGTAAGGTAAAGGACTAAAGTCACCCCTTCACCAATTCGTCCTAGCTACT-----	2640
2592	DB	CGTAAGGTAAA-GACTAAAGTCACCCTTCACCAATTCGTCCTAGCTACTTTGGTTCCTCTCA	2650
2641	QY	-----	2640
2651	DB	GAIAAAAAATTCATGATACTCATTTCTTATGAATCTTTTCAGGAGTTTTTCAGTCCTATT	2710
2641	QY	-----TATTTTTTAAATAATTTCCCTACAAATGATAGCATAAATATGATGAGTGTCTA	2693
2711	DB	AAATTCCTATTTTTTAAATAATTTCCCTACAAATGATAGCATAAATATGATGAGTGTCTA	2770
2694	QY	CACCTTCGTTTTTACTTACTAGTAAATTAATAATAGGAATATCAATATAATGTTTTTA	2753
2771	DB	CACCTTCGTTTTTACTTACTAGTAAATTAATAATAGGAATATCAATATAATGTTTTTA	2829
2754	QY	ATATTTTTTCTTTTCCATATGCTGTAGTCTTACTTAAACTCTGGTGATCCAAACAAAAT	2813
2830	DB	ATATTTTTTCTTTTCCATATGCTGTAGTCTTACTTAAACTCTGGTGATCCAAACAAAAT	2889
2814	QY	GGCTTCAGTGGTGAGATGTCACCTCATGTTATTTCTAGTACTAGAAAACTGAAGACCATG	2873
2890	DB	GGCTTCAGTGGTGAGATGTCACCTCATGTTATTTCTAGTACTAGAAAACTGAAGACCATG	2949











Db 2401 TTAAGTTACAAATGTTTTATGGCCATAAAATTTATTAATAAAAAATGCACAAAACTTTGATTTCT 2460  
Qy 2473 TTTGTATGTAAACAAATTTGTTGTCGTTTTTCAGGCTTTTGTCAATGTCATCTTTTTCATT 2532  
Db 2461 TTTGTATGTAAACAAATTTGTTGTCGTTTTTCAGGCTTTTGTCAATGTCATCTTTTTCATT 2520  
Qy 2533 TTTAAATGTGTTTGTGTTTAAATAGTTAAATAGTTAAATAGTCAAGTTCAAAAATTTCTAAATRTA 2592  
Db 2521 TTTAAATGTGTTTGTGTTTAAATAGTTAAATAGTTAAATAGTCAAGTTCAAAAATTTCTAAATATA 2580  
Qy 2593 CGTAAGGTAAAGACTAAAGTCAACCTTCCACCAATTTGCTAGCTACT 2640  
Db 2581 CGTAAGGTAAA -GACTAAAGTCAACCTTCCACCAATTTGCTAGCTACTTTGGTTCCCTCA 2639  
Qy 2641 ----- 2640  
Db 2640 GAAAAAAATTCATGATACTCATTTCTTATGAATCTTTCCAGGGATTTTTGAGTCTCATTC 2699  
Qy 2641 -----TATTTTTTAAATAATTTCTTACACAAATGATAGATAACATATGCGAGTGTCTA 2693  
Db 2700 AATATCTCTATTTTAAATAATTTCTTACACAAATGATAGATAACATATGCGAGTGTCTA 2759  
Qy 2694 CACCTTGCTTTTACTTAGTAAGATTTAAAAATTTATAGGAATATCAATATAATGTTTTTA 2753  
Db 2760 CACCTTGCTTTTACTTAGT -AGATTAATAATTTATAGGAATATCAATATAATGTTTTTA 2818  
Qy 2754 ATATTTTTTCTTTTCCATATGCTGTAGTCTTACCTAAACTCTGGTGATCCAAACAAAT 2813  
Db 2819 ATATTTTTTCTTTTCCATATGCTGTAGTCTTACCTAAACTCTGGTGATCCAAACAAAT 2878  
Qy 2814 GGCCTCAGTGGTGACAGATGTCACATCATGTTATCTTAGTACTAGAAACTGGAAGACCATG 2873  
Db 2879 GGCCTCAGTGGTGACAGATGTCACATCATGTTATCTTAGTACTAGAAACTGGAAGACCATG 2938  
Qy 2874 TGGAGACTTCATCAAAACATGGGTTTATGTTTCCAGAAATGGAAGACCTGTACCCCTTT 2933  
Db 2939 TGGAGACTTCATCAAAACATGGGTTTATGTTTCCAGCAAGATGGAAGACCTGTACCCCTTT 2998  
Qy 2934 TTTGGTGCTTACTAGCTGGGTGGTGCTGTTTGTAGCTATTTAGAGTCTTAGTGT 2993  
Db 2999 TTTGGTGCTTACTAGCTGGGTGGTGCTGTTTGTAGCTATTTAGAGTCTTAGTGT 3058  
Qy 2994 CTTACTTATAAAGTAGAAATGGTGAGATTTGTTTCTTTTCTACCKTAAAGGGAGATGGT 3053  
Db 3059 CTTACTTATAAAGTAGAAATGGTGAGATTTGTTTCTTTTCTACCTTAAAGGGAGATGGT 3118  
Qy 3054 AAGAAAACAATGAATGTCTTTTTCAAACTTTATGACAGTGAATTTCAAGTCTGTGTTTC 3113  
Db 3119 AAGAAAACAATGAATGTCTTTTTCAAACTTTATGACAAAGTGAATTTTCAAGTCTGTGTTTC 3178  
Qy 3114 AAAAAATATTCATGTACCTGTGATCCAGCAAGAGGGAGTTCCAGTCAAGATCACTAC 3173  
Db 3179 AAAAAATATTCATGTACCTGTGATCCAGCAAGAGGGAGTTCCAGTCAAGATCACTAC 3238  
Qy 3174 AACTGATTAAGTCTTTTAGAGAAATGAGAAATGGAACAGTGAAGGAATCGAGGCCATATTTCC 3233  
Db 3239 AACTGATTAAGTCTTTTAGAGAAATGAGAAATGGAACAGTGAAGGAATCGAGGCCATATTTCC 3298  
Qy 3234 ATGACTTCCCTTGTAAACAGAAACAAAGAGGGACAAGAGGCTGCGCTCTACATCACTC 3293  
Db 3299 ATGACTTCCCTTGTAAACAGAAACAAAGAGGGACAAGAGGCTGCGCTCTACATCACTC 3358  
Qy 3294 TCACCTTCCAAATCTTTGTGGAAGTGCATCTTGTCCAGAACCAAAATTAACCTTACTTCCA 3353  
Db 3359 TCACCTTCCAAATCTTTGTGGAAGTGCATCTTGTCCAGAACCAAAATTAACCTTACTTCCA 3418  
Qy 3354 AGTTCTGGCTGCTTGCAGGTGGAACTCCAGCTCCAAAGGGAGTTAGGGAATGAAGTCTTT 3413  
Db 3419 AGTTCTGGCTGCTTGCAGGTGGAACTCCAGCTCCAAAGGGAGTTAGGGAATGAAGTCTTT 3478  
Qy 3414 TTTTTTAAAGCTTCTCAGCCTTCTAGGGGAACAGAAATTTGGGTGAGCCAAATCTGCAATTT 3473  
Db 3479 TTTTTTAAAGCTTCTCAGCCTTCTAGGGGAACAGAAATTTGGGTGAGCCAAATCTGCAATTT 3538

Qy 3474 CTACTACAGGCATTTGAGACCAGTTAGATTATTGAATAATTATAGAGAGTTATGAACACTT 3533  
Db 3539 CTACTACAGGCATTTGAGACCAGTTAGATTATTGAATAATTATAGAGAGTTATGAACACTT 3598  
Qy 3534 AAATATGATAGTGGTATGACATTTGGATAGAACATGCGGATACCTTTAGAAGTAGAATTGAC 3593  
Db 3599 AAATATGATAGTGGTATGACATTTGGATAGAACATGCGGATACCTTTAGAAGTAGAATTGAC 3658  
Qy 3594 AGGCATATTTAGTTGATGAAATGAGTCAATTTGAGTCTTTTAATAGCCATGATCATTAAT 3653  
Db 3659 AGGCATATTTAGTTGATGAAATGAGTCAATTTGAGTCTCTTAATAGCCATGATCATTAAT 3718  
Qy 3654 TACCAGGTGAAGCTGCTGGAACATATGCTCTCCATTTTACAGTGAAGGAATATAATGAC 3713  
Db 3719 TACCAGGTGAAGCTGCTGGAACATATGCTCTCCATTTTACAGTGAAGGAATATAATGAC 3778  
Qy 3714 AGATTAAATATTTGTTTCTGTCAATCCCAATTTCTTAAGGAAGACTGCGCTACTAT 3773  
Db 3779 AGATTAAATATTTGTTTCTGTCAATCCCAATTTCTTAAGGAAGACTGCGCTACTAT 3838  
Qy 3774 AGCAGTTTTTATTTGTTCAATTTATGAATATAATGAATGAGGAGTCTGTGACTCCTCG 3833  
Db 3839 AGCAGTTTTTATTTGTTCAATTTATGAATATAATGAATGA -GAGTCTGTGACTCCTCG 3897  
Qy 3834 TCTTTACAAAATATTTGGGTGTTGTCAGTATTTTCCCTTTTAAACMTTCCCAATTCGGG 3893  
Db 3898 TCTTTACAAAATATTTGG- - -TGTGTGAGTATTTTCCCTTTTAACCATTCCAATCGGTG 3953  
Qy 3894 TGTGTAGTGGATGTTTCCATTTGGTTTTTAATTTGTATATCCCTGATAGCTATAATGG 3953  
Db 3954 TGTGTAGTGG- - -TTTCATTTTGGTTTTTAATTTGTATATCCCTGATAGCTATAATGG 4008  
Qy 3954 GTCATAGAAATCTTTATACATTTCTAGCAAGTCTCTGTCGAGATATACGTATTGAGA 4013  
Db 4009 GTCATAGAAATCTTTATACATTTCTAGCAAGTCTCTGTCGAGATATACGTATTGAGA 4068  
Qy 4014 TATTACACTAGTCTGTGCTGACTGTTTTCTTTATGCTTTTGTGATGAATAAGAGTTTT 4073  
Db 4069 TATTACACTAGTCTGTGCTGACTGTTTTCTTTATGCTTTTGTGATGAATAAGAGTTTT 4128  
Qy 4074 AAATTTTGACAAAGTCAAAATTTTCTTTTGTGTTGATATTTTCTCTCCAAATTT 4133  
Db 4129 AAATTTTGACAAAGTCAAAATTTTCTTTTGTGTTGATATTTTCTCTCCAAATTT 4188  
Qy 4134 RACCCCAAGNTTCAGATATTTGCTCTATATATAAATTTATATTTTATATTGAGA 4193  
Db 4189 RACCCCAAGNTTCAGATATTTGCTCTATATATAAATTTATATTTTATATTGAGA 4248  
Qy 4194 TCTACCTTGAATTTGATATGTTGTGAATTTATGATCAGGTTCTTTTTTCCCCCAT 4253  
Db 4249 TCTACCTTGAATTTGATATGTTGTGAATTTATGATCAGGTTCTTTTTTCCCCCAT 4308  
Qy 4254 ACAAGTATCCAGTCAATGTAACACTGTTTATTTGAAGAATTAATCTTCTCATTAATTT 4313  
Db 4309 ACAAGTATCCAGTCAATGTAACACTGTTTATTTGAAGAATTAATCTTCTCATTAATTT 4368  
Qy 4314 ACCTTGCCAAATTAGTAAAAA 4333  
Db 4369 ACCTTGCCAAATTAGTAAAAA 4388

## RESULT 4

AR233599 LOCUS AR233599 4386 bp DNA linear PAT 20-DEC-2002  
DEFINITION Sequence 1 from patent US 6458534.  
ACCESSION AR233599  
VERSION AR233599.1 GI:27276201  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 4386)



AUTHORS Concannon, P.J., Vissinga, C.S., Cerosaletti, K.M., Varon-Mateeva, R.,  
Sperling, K. and Reis, A.W.S.  
TITLE Gene associated with NiJmegen breakage syndrome, it's gene product  
and methods for their use  
JOURNAL Patent: US 6458534-A 1 01-OCT-2002;  
FEATURES Location/Qualifiers  
source 1. .4386  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN	Query Match 93.38; Score 4108.4; DB 6; Length 4386; Best Local Similarity 97.19; Pred. No. 0; Matches 4272; Conservative 7; Mismatches 28; Indels 91; Gaps 6;
Qy	28 GCCCCAGCCCTGAGGAGCCGCGATGTGGAACATGCTGCCGCCGCCGCCGCCGCCGCGCAGG 87
Db	1 GCCCCAGCCCTGAGGAGCCGCGATGTGGAACATGCTGCCGCCGCCGCCGCCGCCGCGCAGG 60
Qy	88 AGGAGAACCATACAGACTTTTGACTGGCGTTGAGTACGTTGTTGGAGGAAAAAATGTGC 147
Db	61 AGGAGAACCATACAGACTTTTGACTGGCGTTGAGTACGTTGTTGGAGGAAAAAATGTGC 120
Qy	148 CATTCTTAATTTGAAATGATCAGTCCGATCAGCCGGAATCATGCTGTGTTAACTGCTAACTT 207
Db	121 CATTCTGATTTGAAATGATCAGTCCGATCAGCCGGAATCATGCTGTGTTAACTGCTAACTT 180
Qy	208 TTCTGTAAACCACTGAGTCAAAACAGATGAAATCCCTGTATGACATTTAAAGATAATTC 267
Db	181 TTCTGTAAACCACTGAGTCAAAACAGATGAAATCCCTGTATGACATTTAAAGATAATTC 240
Qy	268 TAAGTATGTPACCTTTGTTAATGAGGAAAAAATGCAGAAATGGCTTTTCCCGAACTTTGAA 327
Db	241 TAAGTATGTPACCTTTGTTAATGAGGAAAAAATGCAGAAATGGCTTTTCCCGAACTTTGAA 300
Qy	328 GTCCGGGGATGTTACTTTTGGAGTGTGTTGAGTGTCTCTGGGAAACCTGTTTAAATCAAGCTAT 447
Db	301 GTCCGGGGATGTTACTTTTGGAGTGTGTTGAGTGTGTTGAGTGTGTTTAAATCAAGCTAT 420
Qy	448 ATTTGCAACTTTGGAGGATTTACTGTAAACAAATTTGGACAGAAAGATGCACTCACCTTTGTCAT 507
Db	421 ATTTGCAACTTTGGAGGATTTACTGTAAACAAATTTGGACAGAAAGATGCACTCACCTTTGTCAT 480
Qy	508 GGTATCAGTGAAAGTTACCAATTTAAACAAATATGTGCATCTATTTGGAGCTCCAATTTGT 567
Db	481 GGTATCAGTGAAAGTTACCAATTTAAACAAATATGTGCATCTATTTGGAGCTCCAATTTGT 540
Qy	568 AAAGCCAGAAATTTTACTGAAATTTCTGAAAGCAGTTTCAGTCCCAAGAGCAGCTCCACA 627
Db	541 AAAGCCAGAAATTTTACTGAAATTTCTGAAAGCAGTTTCAGTCCCAAGAGCAGCTCCACA 600
Qy	628 AATTGAAAGTTTAAACCCACTCTTGATGAACCATCTATTGGAAGTAAAAATGTTGATCT 687
Db	601 AATTGAAAGTTTAAACCCACTCTTGATGAACCATCTATTGGAAGTAAAAATGTTGATCT 660
Qy	688 GTCAGAGCGCAGGAAGAAAAAACAATCTTTCAAAGGAAAAACAATTTATTTTGAATGC 747
Db	661 GTCAGAGCGCAGGAAGAAAAAACAATCTTTCAAAGGAAAAACAATTTATTTTGAATGC 720
Qy	748 CAAACAGCAATAAGAAATTTGAGTTCCGAGTTGCTTTGGAGGTGGGAGAGCTAGGTTGAT 807
Db	721 CAAACAGCAATAAGAAATTTGAGTTCCGAGTTGCTTTGGAGGTGGGAGAGCTAGGTTGAT 780
Qy	808 AACAGAAGAAATGAAGAGAAACAATTTCTTTTGGCTCCGCGGAACTGTGTTGTTGA 867
Db	781 AACAGAAGAAATGAAGAGAAACAATTTCTTTTGGCTCCGCGGAACTGTGTTGTTGA 840
Qy	868 TACAGGAATAAACAATCTCAGACCTTTAATCTCTGATGTGCAGAGAAATGGATTCAGTC 927

Db	841 TACAGGAATAAACAATCTCAGACCTTAATTCCTGACTGTGCAGAGAAATGGATTCAGTC 900
Qy	928 AATAATGGATATGCTCCAAAGGCAAGGTCTTAGACCTATTCTTGAAAGCAGAAATTTGGATT 987
Db	901 AATAATGGATATGCTCCAAAGGCAAGGTCTTAGACCTATTCTTGAAAGCAGAAATTTGGATT 960
Qy	988 GCGGGTGAATTTTCATGACTACAAAGAAATTAATCTGTGATTCCTCAGGGCCATCCCAAGTACAGG 1047
Db	961 GCGGGTGAATTTTCATGACTACAAAGAAATTAATCTGTGATTCCTCAGGGCCATCCCAAGTACAGG 1020
Qy	1048 ATTAAGACAAACAATCTCCAGGACCAAGCCTTTTCAAGGCGGTGTGAGTTGATCAAAAATCT 1107
Db	1021 ATTAAGACAAACAATCTCCAGGACCAAGCCTTTTCAAGGCGGTGTGAGTTGATCAAAAATCT 1080
Qy	1108 AATGCCAAGCGCCGCCAGTGAACAATCAACATACATACATACATACATACATACATACATACAT 1167
Db	1081 AATGCCAAGCGCCGCCAGTGAACAATCAACATACATACATACATACATACATACATACATACAT 1140
Qy	1168 AGATACATGGGATTTGAGTGAAGGCAAAAGAAATCAAAAGTCTCCAAATGGAACAAAA 1227
Db	1141 AGATACATGGGATTTGAGTGAAGGCAAAAGAAATCAAAAGTCTCCAAATGGAACAAAA 1200
Qy	1228 ATTCAGAAATGCTTTTCAACAGACGACCCACTGTAAAGGAGTCTCTGCAAAACAGCTCTAA 1287
Db	1201 ATTCAGAAATGCTTTTCAACAGATGCACCCACTGTAAAGGAGTCTCTGCAAAACAGCTCTAA 1260
Qy	1288 TAATAATAGTATGGTATCAATTAATTTGGCTAAGATGAGAAATCCCAAACTATCAGCTTTTC 1347
Db	1261 TAATAATAGTATGGTATCAATTAATTTGGCTAAGATGAGAAATCCCAAACTATCAGCTTTTC 1320
Qy	1348 ACCAATCTAAATGCGCAAGTATAAATAAAGTAAAGATAGGGCTTTCTCAGCAGCAGCAGAC 1407
Db	1321 ACCAATCTAAATGCGCAAGTATAAATAAAGTAAAGATAGGGCTTTCTCAGCAGCAGCAGAC 1380
Qy	1408 CAACTCCATCAGAAATCTTTTCCAGCCCTTACCAAAAAAGGAAAGGATGAAGAAA 1467
Db	1381 CAACTCCATCAGAAATCTTTTCCAGCCCTTACCAAAAAAGGAAAGGATGAAGAAA 1440
Qy	1468 TCAAGAAATGTCTTCATGCAAAATCAGCAAGATAGAAAAGTCTTTCTTTTGAACA 1527
Db	1441 TCAAGAAATGTCTTCATGCAAAATCAGCAAGATAGAAAAGTCTTTCTTTTGAACA 1500
Qy	1528 AACACAATCTGTACACCTCAATTTGGGAAAAAATAAGGAGCAGCATCTATCTGAGAAATGA 1587
Db	1501 AACACAATCTGTACACCTCAATTTGGGAAAAAATAAGGAGCAGCATCTATCTGAGAAATGA 1560
Qy	1588 GCTGTGGACACAAATCTCAGCAATTAATTTTACAGATCAGATTTAAATCTATTCT 1647
Db	1561 GCTGTGGACACAAATCTCAGCAATTAATTTTACAGATCAGATTTAAATCTATTCT 1620
Qy	1648 GAAAAATCTGCCAGTAAATCTCATGCTGCAGAAAGCTTAAGATCAAAATAAAAAAGGGA 1707
Db	1621 GAAAAATCTGCCAGTAAATCTCATGCTGCAGAAAGCTTAAGATCAAAATAAAAAAGGGA 1680
Qy	1708 AATGATGATGTGGCCATAGAAAGATGAAGTATTTGGAACAGTTATTCAAGGACACAAAAAC 1767
Db	1681 AATGATGATGTGGCCATAGAAAGATGAAGTATTTGGAACAGTTATTCAAGGACACAAAAAC 1740
Qy	1768 AGAGTTAGAAATTTGATGTGAAAGTTCAAAAAAGGAGGAGATGTCAATGTTAGAAAAAG 1827
Db	1741 AGAGTTAGAAATTTGATGTGAAAGTTCAAAAAAGGAGGAGATGTCAATGTTAGAAAAAG 1800
Qy	1828 GCCAAGGATGGATATAGAAAAAATAATGACATTTTTCAGTGTATGAAGCAGTACCAAGAAAGTAG 1887
Db	1801 GCCAAGGATGGATATAGAAAAAATAATGACATTTTTCAGTGTATGAAGCAGTACCAAGAAAGTAG 1860
Qy	1888 CAAAAATCTCAAGAAAAATGAAATTTGGGAGAAAAAGTGAATCTCAAGGAAAGCTCACTATG 1947
Db	1861 CAAAAATCTCAAGAAAAATGAAATTTGGGAGAAAAAGTGAATCTCAAGGAAAGCTCACTATG 1920
Qy	1948 GTCAGCTAAAGAAATATCTTAAACATGACAACTTCCAGGATGATAGTGAGATGCTTCCAAA 2007
Db	1921 GTCAGCTAAAGAAATATCTTAAACATGACAACTTCCAGGATGATAGTGAGATGCTTCCAAA 1980



QY 2008 AAGCTGTTATTCACCTGAATTTTAGATCCTCGTGATTTAATAAATCTTACTTCCAGAAATCC 2067  
DB 1981 AAGCTGTTATTCACCTGAATTTTAGATCCTCGTGATTTAATAAATCTTACTTCCAGAAATCC 2040  
QY 2068 GTCTGGCATAAATGATGATTTATGGTCAACTTAAAAATTTCAAGAAATTTCAAAAAGGTCAC 2127  
DB 2041 ATCTGGCATAAATGATGATTTATGGTCAACTTAAAAATTTCAAGAAATTTCAAAAAGGTCAC 2100  
QY 2128 ATATCTGGAGCAGGAAAACTTCCACACATCATTTGGAGGATCAGATCTAATAGCTCATCA 2187  
DB 2101 ATATCTGGAGCAGGAAAACTTCCACACATCATTTGGAGGATCAGATCTAATAGCTCATCA 2160  
QY 2188 TGCTCGAAGAAATACAGACTAGAGAGTGGCTAAGGCAGGAAATGGAGTCAAAATCA 2247  
DB 2161 TGCTCGAAGAAATACAGAACTAGAGAGTGGCTAAGGCAGGAAATGGAGTCAAAATCA 2220  
QY 2248 ACATGCAAAAGAAAGAGTCTCTTCTGATGATCTTTTATAGATACAATCCTTATTTAAAAAG 2307  
DB 2221 ACATGCAAAAGAAAGAGTCTCTTCTGATGATCTTTTATAGATACAATCCTTATTTAAAAAG 2280  
QY 2308 GAGAAGATAACTGAGGATTTTAAAAAGAGCCATGGAAAAAATCTTCTAGTAAGCATCTAC 2367  
DB 2281 GAGAAGATAACTGAGGATTTTAAAAAGAGCCATGGAAAAAATCTTCTAGTAAGCATCTAC 2340  
QY 2368 TTCAGGCCAACAGGTTTATATGAATATATAGTGTATAGCGGATTTAAGTTACATGTT 2427  
DB 2341 TTCAGGCCAACAGGTTTATATGAATATATAGTGTATAGCGGATTTAAGTTACATGTT 2400  
QY 2428 TTATGGCCTAAATTTTAAATAAAATGCACAAAATCTTTGATCTTTTGTATGTAAACAAT 2487  
DB 2401 TTATGGCCTAAATTTTAAATAAAATGCACAAAATCTTTGATCTTTTGTATGTAAACAAT 2460  
QY 2488 TGTTGTGTCGTTTTGAGGCTTTGTCATTGTCATCTTTTTCATTTTAAATGTTGTTTG 2547  
DB 2461 TGTTGTGTCGTTTTGAGGCTTTGTCATTGTCATCTTTTTCATTTTAAATGTTGTTTG 2520  
QY 2548 TTATTAATAATAGTTAATATAGTCACAGTTCAAAATTTCTAAATTTAGTAAGTAAGGAC 2607  
DB 2521 TTATTAATAATAGTTAATATAGTCACAGTTCAAAATTTCTAAATTTAGTAAGGTAAGGAC 2579  
QY 2608 TAAAGTCACCTTCCACCAATTTGCTAGTACT- 2640  
DB 2580 TAAAGTCACCTTCCACCAATTTGCTAGTACTTGGTCCCTCAGAAAAAATTCATGA 2639  
QY 2641 -----TATTTTAA 2648  
DB 2640 TACTCATTTTCTTATGAATCTTTCCAGGGAATTTTGTAGTCCTATTTCAAAATTCCTATTTTAA 2699  
QY 2649 AATAATTTCTACACAAATGATAGCATAAATATGATGAGTGTCTACACCTTGTCTTTTAA 2708  
DB 2700 AATAATTTCTACACAAATGATAGCATAAATATGATGAGTGTCTACACCTTGTCTTTTAA 2759  
QY 2709 CTTAGTAAGATTTAAAAATTTATAGGAATATCAATATAATGTTTTTAAATTTTTTCTTTTC 2768  
DB 2760 CTTAGT-AGATTTAAATTTATAGGAATATCAATATAATGTTTTTAAATTTTTTCTTTTC 2818  
QY 2769 CATATGCTGTAGTCTTACCTAACTCTGCTGTGATCCAAAACAAAATGGCTTCAGTGTGCA 2828  
DB 2819 CATATGCTGTAGTCTTACCTAACTCTGCTGTGATCCAAAACAAAATGGCTTCAGTGTGCA 2878  
QY 2829 GATGTCACCTACATGTTATCTAGTACTAGAACTGAAGCAAGCAATGGAGACTTCATCAA 2888  
DB 2879 GATGTCACCTACATGTTATCTAGTACTAGAACTGAAGCAAGCAATGGAGACTTCATCAA 2938  
QY 2889 ACATGGGTTTGTGTTTTACACAGAAATGGAAGACCTGTACCCCTTTTGTGGTCTTACTG 2948  
DB 2939 ACATGGGTTTGTGTTTTACACAGAAATGGAAGACCTGTACCCCTTTTGTGGTCTTACTG 2998  
QY 2949 AGCTGGGTTGGTGTCTGTTTTGAGCTTATTTAGAGTCCTAGTTTTTCTCTTATAAAGTA 3008  
DB 2999 AGCTGGGTTGGTGTCTGTTTTGAGCTTATTTAGAGTCCTAGTTTTTCTCTTATAAAGTA 3058

QY 3009 GAAATGGTGAGATTTGTTTTCTTTTCTACCKTAAAGGAGATGGTAAGAAAACAATGAATG 3068  
DB 3059 GAAATGGTGAGATTTGTTTTCTTTTCTACCTTAAAGGAGATGGTAAGAAAACAATGAATG 3118  
QY 3069 TCTTTTTTCAAACTTTTATTTGACAAAGTGAATTTTCAAGTCTGTGTTTCAAAAAATATATTCATG 3128  
DB 3119 TCTTTTTTCAAACTTTTATTTGACAAAGTGAATTTTCAAGTCTGTGTTTCAAAAAATATATTCATG 3178  
QY 3129 TACCTGTGATCCAGCAAGAGGAGTTCCAGTCAAGAGTCACCTAACAATGATAGTTGTT 3188  
DB 3179 TACCTGTGATCCAGCAAGAGGAGTTCCAGTCAAGAGTCACCTAACAATGATAGTTGTT 3238  
QY 3189 TAGAAGATGAGAAATGGAACAGTGAAGGATGAGGCCATATTTTCATGACTTCCCTTGTGA 3248  
DB 3239 TAGAAGATGAGAAATGGAACAGTGAAGGATGAGGCCATATTTTCATGACTTCCCTTGTGA 3298  
QY 3249 AACAGAAGCAACAGAAAGGACAAAGAGGTGGCTCTCATCATCTCTCACCTTCCAAATCT 3308  
DB 3299 AACAGAAGCAACAGAAAGGACAAAGAGGTGGCTCTCATCATCTCTCACCTTCCAAATCT 3358  
QY 3309 TGTGAAAGTGCATCTACTTGCAGAACCAAAATTAATCTTACTTCCAGTTCTGGCTGTG 3368  
DB 3359 TGTGAAAGTGCATCTACTTGCAGAACCAAAATTAATCTTACTTCCAGTTCTGGCTGTG 3418  
QY 3369 CAGGTGGAACTCCAGCTCAAGGAGTTAGGGAATGAAGTCTTTTTTAAAGCTTCT 3428  
DB 3419 CAGGTGGAACTCCAGCTCAAGGAGTTAGGGAATGAAGTCTTTTTTAAAGCTTCT 3478  
QY 3429 CAGCCTTCTAGGGAACAGAAATTTGGTGAGCCAATCTGCAATTTCTACTACAGGCAATG 3488  
DB 3479 CAGCCTTCTAGGGAACAGAAATTTGGTGAGCCAATCTGCAATTTCTACTACAGGCAATG 3538  
QY 3489 AGACAGGTAGATTATTAATTAATATATAGAGATTATGAACACTTAAATATATGATAGTGG 3548  
DB 3539 AGACAGGTAGATTATTAATTAATATATAGAGATTATGAACACTTAAATATATGATAGTGG 3598  
QY 3549 TATGACATTTGGATAGAACATGGGATCTTTTAAAGTAGAATTTGACAGGCAATATAGTTG 3608  
DB 3599 TATGACATTTGGATAGAACATGGGATCTTTTAAAGTAGAATTTGACAGGCAATATAGTTG 3658  
QY 3609 ATGAAATGGAGTCATTTGAGTCTTTTAAATAGCCATGTATCATTAATTAACAAGTGAAGCTG 3668  
DB 3659 ATGAAATGGAGTCATTTGAGTCTTTTAAATAGCCATGTATCATTAATTAACAAGTGAAGCTG 3718  
QY 3669 GTGGAACATATGCTTCCATTTTACAGTTTAGGAATATAAGGACAGATTAATATATGTTT 3728  
DB 3719 GTGGAACATATGCTTCCATTTTACAGTTTAGGAATATAAGGACAGATTAATATATGTTT 3778  
QY 3729 TCTGTCAATGCCCCAACATCCCTTTCTAAAGGAAGACTGCCCTACTATAGCAGTTTATATAT 3788  
DB 3779 TCTGTCAATGCCCCAACATCCCTTTCTAAAGGAAGACTGCCCTACTATAGCAGTTTATATAT 3838  
QY 3789 TGTCAATTTATGAATATAATGAATGAGGTTCTGGTACCTCTCTCTTTTACAAATATG 3848  
DB 3839 TGTCAATTTATGAATATAATGAATGA-GAGTTCTGGTACCTCTCTCTTTTACAAATATG 3897  
QY 3849 GGTGTTGTCAGTATTTTTTCCCTTTTAAACCMTTTCCCAATTCGGGTGTAGTGGATGT 3908  
DB 3898 G-----TGTGTCAGTATTTTTTCCCTTTTAAACCMTTTCCCAATTCGGGTGTAGTGGATGT 3948  
QY 3909 TTCCATTTTGGGTTTTAAATTTGTATATCCCTGATAGTATTAATTTGGGTATAGAAATTTCTT 3968  
DB 3949 TTTCAATTTTGGTTTTAAATTTGTATATCCCTGATAGTATTAATTTGGGTATAGAAATTTCTT 4008  
QY 3969 TATACATCTTAGATCAAGTCTCTTGYCGGATATACGTTATGAGATATTAACACCTAGTCT 4028  
DB 4009 TATACATCTTAGATCAAGTCTCTTGYCGGATATATGTTATGAGATATTAACACCTAGTCT 4068  
QY 4029 GTGGCTTGCATGTTTTCTTTTATGCTTTTGTAGTAAGAGTTTAAATTTTGAAGGT 4088  
DB 4069 GTGGCTTGCATGTTTTCTTTTATGCTTTTGTAGTAAGAGTTTAAATTTTGAAGGT 4128  
QY 4089 CAAATTTATTTTTTTTCTTTTGTGATATTTTTTCTCTCCAATTTTAAACCCCAAGATTTTCA 4148







QY	433	TTTTAAATCAAGCTATATTGCAACTTGGAGGATTTTACTGTAAACAATTTGGACGAGAAATG	492	QY	1512	GTTCCTCTTTTAGAACAAACACAACTCTGTACACCTCTATTGTGGAATAATAAGGACGAC	1571
Db	487	TTTTAAATCAAGCTATATTGCAACTTGGAGGATTTTACTGTAAACAATTTGGACGAGAAATG	546	Db	1567	GTTCCTCTTTTAGAACAAACACAACTCTGTACACCTCTATTGTGGAATAATAAGGACGAC	1626
QY	493	CACCTACCTTTGTTCATGGTATCAGTGAAGATTACCATTTAAACAATATGTGCACCTCATTTTG	552	QY	1572	ATCTATCTGAGAAATGAGCTGTGGACAACTCTAGACAATAAATCTTATTTTACAGATACAG	1631
Db	547	CACCTACCTTTGTTCATGGTATCAGTGAAGATTACCATTTAAACAATATGTGCACCTCATTTTG	606	Db	1627	ATCTATCTGAGAAATGAGCTGTGGACAACTCTAGACAATAAATCTTATTTTACAGATACAG	1686
QY	553	TGAGCTGCCAATTTGTAAGCCAGAAATATTCTGAATTCCTGAAAGCAGTTTCAGTCCAA	612	QY	1632	ATTTAAATCTTATCTGAAATTTCTGCCAGTAAATCTCATCTGCAGAAAGCTTAAGAT	1691
Db	607	TGAGCTGCCAATTTGTAAGCCAGAAATATTCTGAATTCCTGAAAGCAGTTTCAGTCCAA	666	Db	1687	ATTTAAATCTTATCTGAAATTTCTGCCAGTAAATCTCATCTGCAGAAAGCTTAAGAT	1746
QY	613	GAGACGCTCCCAAAATGAAAGTTTACCCACCTCTTGATGAACCATCTATTTCGAAG	672	QY	1692	CAATTAATAAAGGAAATGGAATGATGTGGCCATAGAAGTGAAGTATTGGAAACAGTTAT	1751
Db	667	GAGACGCTCCCAAAATGAAAGTTTACCCACCTCTTGATGAACCATCTATTTCGAAG	726	Db	1747	CAATTAATAAAGGAAATGGAATGATGTGGCCATAGAAGTGAAGTATTGGAAACAGTTAT	1806
QY	673	TAAAAATGTTGATCTGTCCAGGACGGCAGGAAAGAAACAAATCTTCAAAAGGAAACATTT	732	QY	1752	TCAAGGACACAAACACAGAGTTAGAAAATTGATGTGAAAATTTCAAAAAACAGGAGGAATG	1811
Db	727	TAAAAATGTTGATCTGTCCAGGACGGCAGGAAAGAAACAAATCTTCAAAAGGAAACATTT	786	Db	1807	TCAAGGACACAAACACAGAGTTAGAAAATTGATGTGAAAATTTCAAAAAACAGGAGGAATG	1866
QY	733	TATATTTTGAATGCCCACAGCATTAAGAAATTTGAGTTCCGCAAGTTGTCTTTGGAGGTGG	792	QY	1812	TCAATGTTAGAAAAGGCCAGGATGGATATAGAAACAAATGACACTTTTCAGTGATGAAG	1871
Db	787	TATATTTTGAATGCCCACAGCATTAAGAAATTTGAGTTCCGCAAGTTGTCTTTGGAGGTGG	846	Db	1867	TCAATGTTAGAAAAGGCCAGGATGGATATAGAAACAAATGACACTTTTCAGTGATGAAG	1926
QY	793	GGAAGCTAGTGTGATACAGAGAAATGAAGAAGAACATATTCTTTTGGCTCCGGG	852	QY	1872	CAGTACCAGAAAGTAGCAAAATATCTCAAGAAAATGAAATTTGGGAAGAAACCTGAACTCA	1931
Db	847	GGAAGCTAGTGTGATACAGAGAAATGAAGAAGAACATATTCTTTTGGCTCCGGG	906	Db	1927	CAGTACCAGAAAGTAGCAAAATATCTCAAGAAAATGAAATTTGGGAAGAAACCTGAACTCA	1986
QY	853	AACGTGTGTTGTGATACAGGAAATAACAACTCACAGACCTTAAATTCCTGACTGTGAGAA	912	QY	1932	AGGAAGACTACTATGGTCAGCTAAAGAAAATCTTAAACAATGACAAAATTTAGAGATGATA	1991
Db	907	AACGTGTGTTGTGATACAGGAAATAACAACTCACAGACCTTAAATTCCTGACTGTGAGAA	966	Db	1987	AGGAAGACTACTATGGTCAGCTAAAGAAAATCTTAAACAATGACAAAATTTAGAGATGATA	2046
QY	913	GAAATGGAATTCAGTCAATATGGAATGCTCCAAAGGCAAGTCTTAGACCTATTCTCTGA	972	QY	1992	GTGAGATGCTTCCAAAAGAGCTGTTATTGACCTGAATTTAGATCACTGGTGATTTAAAACT	2051
Db	967	GAAATGGAATTCAGTCAATATGGAATGCTCCAAAGGCAAGTCTTAGACCTATTCTCTGA	1026	Db	2047	GTGAGATGCTTCCAAAAGAGCTGTTATTGACCTGAATTTAGATCACTGGTGATTTAAAACT	2106
QY	973	AGCAGAAATTTGGATTCGGCGTGATTTTCATGACTACAAAGAAATTAATCTGTGATCCTCAGG	1032	QY	2052	CTACTTCCAGAAATCCGTCTGGCATAAATGATGATTTATGGTCAACTAAAAAATTTCAAGA	2111
Db	1027	AGCAGAAATTTGGATTCGGCGTGATTTTCATGACTACAAAGAAATTAATCTGTGATCCTCAGG	1086	Db	2107	CTACTTCCAGAAATCCATCTGGCATAAATGATGATTTATGGTCAACTAAAAAATTTCAAGA	2166
QY	1033	CCATCCAGTACAGGATTTAAAGACAACTCCAGGACCAAGCTTTTCAAGAGCGGTGTC	1092	QY	2112	AATTCAAAAAGGTGCATATCTCTGGAGCAGGAAAACTTCCACACATCATTTGGAGGATCAG	2171
Db	1087	CCATCCAGTACAGGATTTAAAGACAACTCCAGGACCAAGCTTTTCAAGAGCGGTGTC	1146	Db	2167	AATTCAAAAAGGTGCATATCTCTGGAGCAGGAAAACTTCCACACATCATTTGGAGGATCAG	2226
QY	1093	AGTTGATGAAAACTTAATGCCAAAGCGCCAGTGGAACATAAACAATACGTAGCTGACAC	1152	QY	2172	ATCTAATAGCTCATCATCTCGAAGAAATACAGAACTAGAGAGTGGCTTAAGCGAGGAAA	2231
Db	1147	AGTTGATGAAAACTTAATGCCAAAGCGCCAGTGGAACATAAACAATACGTAGCTGACAC	1206	Db	2227	ATCTAATAGCTCATCATCTCGAAGAAATACAGAACTAGAGAGTGGCTTAAGCGAGGAAA	2286
QY	1153	AGAAATCAGACGACAGATACATGGGATTTGAGTGAAGGCCAAAGAAATCAAACTCTC	1212	QY	2232	TGGAGGTACAAAATCAACATGCAAAAGAGAGTCTCTTGTCTGATGATCTTTTTAGATACA	2291
Db	1207	AGAAATCAGACGACAGATACATGGGATTTGAGTGAAGGCCAAAGAAATCAAACTCTC	1266	Db	2287	TGGAGGTACAAAATCAACATGCAAAAGAGAGTCTCTTGTCTGATGATCTTTTTAGATACA	2346
QY	1213	CAAAATGGAACAAAAATTCAGAAATGCTTTTCAAGACGCAACCCACTGTAAGAGAGTCTTG	1272	QY	2292	ATCTCTTAATTTAAAAAGGAGAGATACTGAGAGATTTTAAAAAGAGCCATGGAATAAATCTT	2351
Db	1267	CAAAATGGAACAAAAATTCAGAAATGCTTTTCAAGACGCAACCCACTGTAAGAGAGTCTTG	1326	Db	2347	ATCTCTTAATTTAAAAAGGAGAGATACTGAGAGATTTTAAAAAGAGCCATGGAATAAATCTT	2406
QY	1273	CAAAAACAGCTCTAATAATAATAGTATGGTATCAAAATACTTTGGCTTAAGATGAGAAATCCC	1332	QY	2352	CCTAGTAGCATCTACTTTCAGGCCCAACAGGTTTATTAATATATAGTATAGTAGAGCGA	2411
Db	1327	CAAAAACAGCTCTAATAATAATAGTATGGTATCAAAATACTTTGGCTTAAGATGAGAAATCCC	1386	Db	2407	CCTAGTAGCATCTACTTTCAGGCCCAACAGGTTTATTAATATATAGTATAGTAGAGCGA	2466
QY	1333	AAACTATCAGCTTTTCAACCACTAAATTTGCCAGTATATAATAAGTAAAGATAGGCGCTTC	1392	QY	2412	TTTAAAGTTACAAATGTTTATGGCTTAAATTTTAAATAAATAATGACAAAACTTTGATTTC	2471
Db	1387	AAACTATCAGCTTTTCAACCACTAAATTTGCCAGTATATAATAAGTAAAGATAGGCGCTTC	1446	Db	2467	TTTAAAGTTACAAATGTTTATGGCTTAAATTTTAAATAAATAATGACAAAACTTTGATTTC	2526
QY	1393	TCAGACGACGACAGCAACTCCATCAGAAACTACTTTTCAGCCGCTCTACCC-AAAAAAGGG	1451	QY	2472	TTTTCTATGTAACAATGTTTGTCTGTTTTCAGGCTTTGTCATTTGCAATCTTTTTTTTCAT	2531
Db	1447	TCAGACGACGACAGCAACTCCATCAGAACTACTTTTCAGCCGCTCTACCC-AAAAAAGGG	1506	Db	2527	TTTTCTATGTAACAATGTTTGTCTGTTTTCAGGCTTTGTCATTTGCAATCTTTTTTTTCAT	2586
QY	1452	AAAGGATGAGAAAAATCAAGAAATGTTCTGTCGAAATCAGCAAGATAGAAACGCTCTT	1511	QY	2532	TTTTTAAATGCTGTTTGTGTTTATTAATATAGTATAGTACACAGTTTCAAAATTTCTAAATRT	2591
Db	1507	AAAGGATGAGAAAAATCAAGAAATGTTCTGTCGAAATCAGCAAGATAGAAACGCTCTT	1566	Db	2587	TTTTTAAATGCTGTTTGTGTTTATTAATATAGTATAGTACACAGTTTCAAAATTTCTAAATRT	2646
				QY	2592	ACGTAAGGTAAAGGACTAAAGTCAACCTTCCACCATTGTCTTACTACT- - - - -	2640



[illegible]

Db	3725	CAGGGCATAATTAGTTGATGAATAATGGAGTCATTTGAGTCTCTTTAATAGCCAGTATCATATAA	3784
Qy	3653	TTACCAAGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGGAATAATAATGGA	3712
Db	3785	TTACCAAGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGGAATAATAATGGA	3844
Qy	3713	CAGATAATATGTGTTCTGTCATGGCCCAAAATCCCTTTCTTAAGGAAGACTGCCCTACTA	3772
Db	3845	CAGATTAAATTTGTTCTCTGTCTGATGCCCAAAATCCCTTTCTTAAGGAAGACTGCCCTACTA	3904
Qy	3773	TAGCAGTTTTTATATTTGTTCAATTTATGAATAATAAGAAATGAGGAGTTCGTGTACCTCCT	3832
Db	3805	TAGCAGTTTTTATATTTGTTCAATTTATGAATAATAAGAAATGAGTTCGTGTACCTCCT	3963
Qy	3833	GTCTTTACAAAATATTTGGGTGTGTGCCAGTATTTTTCCCTTTTAAACOMITCCCAATTCGG	3892
Db	3964	GTCTTTACAAAATATTTGG---TGTTGTCAAGTATTTTTCCCTTTTAAACATTCCAATCCGGT	4019
Qy	3893	GTGTGTAGGTGGATGTTTTCCATTTGGGTTTTTAATTTGTATATCCCTGATAGCTATAATTC	3952
Db	4020	GTGTAGTGTATG---TTTCATTTTGGTTTTAATTTTGTATATCCCTGATAGCTATAATTC	4074
Qy	3953	GGTCATAGAAAATCTTTATATACATCTAGATGCAAGTCTCTGTGCGGATATACGTATTGAG	4012
Db	4075	GGTCATAGAAAATCTTTATATACATCTAGATGCAAGTCTCTGTGCGGATATATGTATTGAG	4134
Qy	4013	ATATTACACCTAGTCTGTGGCTTGACGTGTTTTCTTTATGTCTTTTGTAGATGAATAGAACTTT	4072
Db	4135	ATATTACACCTAGTCTGTGGCTTGACGTGTTTTCTTTATGTCTTTTGTAGATGAATAGAACTTT	4194
Qy	4073	TAAATTTTGACAAGGTCAAAATTTATTTTTTCTTTTGTGTTGATATTTTTTCTCTCCAAAT	4132
Db	4195	TAAATTTTGACAAGGTCAAAATTTATTTTTTCTTTTGTGTTGATATTTTTTCTCTCCAAAT	4254
Qy	4133	TAAACCCCAAGATTTTCAGATATTCGTCTATATATAAACCCTTATATTTTATATTGTTG	4192
Db	4255	TAAACCCCAAGATTTTCAGATATTCGTCTATATATAAACCCTTATATTTTATATTGTTG	4314
Qy	4193	ATCTACCTTGAATTTGATATGTATGTTGTGAATTTATCGATCAGGGTCTTTTTTCCCCCA	4252
Db	4315	ATCTACCTTGAATTTGATATGTATGTTGTGAATTTATCGATCAGGGTCTTTTTTCCCCCA	4374
Qy	4253	TACAAGTATCCAGTCAATTTGTAACACTGTTTATTTGAAAGAAATATCCTTTCTCATTAAT	4312
Db	4375	TACAAGTATCCAGTCAATTTGTAACACTGTTTATTTGAAAGAAATATCCTTTCTCATTAAT	4434
Qy	4313	TACCTTGCCAAATTTAGTAAAAAAATCAATTA	4342
Db	4435	TACCTTGCCAAATTTAGTAAAAAAATCAATTA	4464

[illegible]



3595 John Hopkins Court, San Diego, CA 92121, USA  
Tel: 18582029018  
Fax: 18582029020  
Email: abraun@sequenom.com  
Primer A: No primer sequence submitted  
Primer B: No primer sequence submitted  
STS size: 4423.

FEATURES

source  
1..4423  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone\_lib="Human DNA (Sequenom)"  
<1..>4423

STS

ORIGIN

Query Match 91.7%; Score 4038.4; DB 11; Length 4423;  
Best Local Similarity 96.6%; Pred. No. 0;  
Matches 4270; Conservative 8; Mismatches 41; Indels 103; Gaps 11;

Qy	12	GGCGGTTGCAGCTCGGCCCGCCAGCCCTGAGGAGCGGACCGGATGTGGAACTGCTGCCCG	71
Db	4406	GGCGGTTGCAGCTCGGCCCGCC--AGCCTGAGGAGCGGACCGGATGTGGAACTGCTGCCCG	4349
Qy	72	CGCGGGCCCGGAGGAGGAGAACCATACAGACTTTTGACTGGCGTTGAGTACGTTGTTG	131
Db	4348	CGGAGCG-CCGGCAGGAGGAGAACCATACAGACTTTTGACTGGCGTTGAGTACGTTGTTG	4290
Qy	132	GAGGAAAACTGTGCCATTCTTAATTGAAAAATGATCAGTCGATCAGCCGAAATCATGCTG	191
Db	4289	GAGGAAAACTGTGCCATTCTTAATTG-ACATGATCAGTCGATCAGCCGAAATCATGCTG	4231
Qy	192	TGTTAACTGCTAACTTTCTGTAAACCAACTGTGATCAACACAGATGAATCCCTGTTTCA	251
Db	4230	TGTTAACTGCTAACTTTCTGTAAACCAACTGTGATCAACACAGATGAATCCCTGTTTCA	4171
Qy	252	CATTAAGAATTAATCTTAAGTATGTTACCTTTGTTAATGAGGAAAAATGCAGATGGCT	311
Db	4170	CATTAAGAATTAATCTTAAGTATGTTACCTTTGTTAATGAGGAAAAATGCAGATGGCT	4111
Qy	312	TTTCCGAACTTTGAAGTCGGGGATGGTATTACTTTTGGAGTGTGGAAGTAAATTC	371
Db	4110	TTTCCGAACTTTGAAGTCGGGGATGGTATTACTTTTGGAGTGTGGAAGTAAATTC	4051
Qy	372	GATAGAGTATGAGCCTTGTGTGATGCTCTCTTTGATGTCCTCTGGGAAACTG	431
Db	4050	GATAGAGTATGAGCCTTGTGTGATGCTCTCTTTGATGTCCTCTGGGAAACTG	3991
Qy	432	CTTTAAATCAAGCTATATTGCAACTTGGAGGATTTACTGTAAACAAATGGACAGAGAAT	491
Db	3990	CTTTAAATCAAGCTATATTGCAACTTGGAGGATTTACTGTAAACAAATGGACAGAGAAT	3931
Qy	492	GCACTCACCTTGTGATGATCAGTGAAAGTTACCAATTAACAAATATGTGCACTCATTT	551
Db	3930	GCACTCACCTTGTGATGATCAGTGAAAGTTACCAATTAACAAATATGTGCACTCATTT	3871
Qy	552	GTGGAGTCCCAATTTAAGCCAGATATTTTACTGAAATTCCTGGAAGCAGTTCACTCCA	611
Db	3870	GTGGAGTCCCAATTTAAGCCAGATATTTTACTGAAATTCCTGGAAGCAGTTGATCCA	3811
Qy	612	AGAAGCAGCCTCCACAAATTTGAAAGTTTTCACCACTCTTGATGAACCATCTATTGGAA	671
Db	3810	AGAAGCAGCCTCCACAAATTTGAAAGTTTTCACCACTCTTGATGAACCATCTATTGGAA	3751
Qy	672	GTAATAATGTTGATCTGTGAGGACGGCAGGAAAGAAAAAACAATCTTCAAGGGAAAAAT	731
Db	3750	GTAATAATGTTGATCTGTGAGGACGGCAGGAAAGAAAAAACAATCTTCAAGGGAAAAAT	3691
Qy	732	TTATATTTTGTGATGCCAACACATAGAAATTTGAGTTCCGAGTTGCTTTTGGAGGTG	791
Db	3690	TTATATTTTGTGATGCCAACACATAGAAATTTGAGTTCCGAGTTGCTTTTGGAGGTG	3631
Qy	792	GGGAAGCTAGTTGTATACAGAGAGATGAAGAGAACATAATTTCTTTTGGCTCCCG	851

Db	3630	GGGAAGCTAGTTGTATACAGAGAGATGAAGAGAACATAATTTCTTTTGGCTCCCG	3571
Qy	852	GAACTGTGTTGTTGATACAGGAATAACAAACTACAGACCTTAATTTCTGACTGTCAGA	911
Db	3570	GAACTGTGTTGTTGATACAGGAATAACAAACTACAGACCTTAATTTCTGACTGTCAGA	3511
Qy	912	AGAAATGGAATTCAGTCAATTAATGGATATGCTCCAAAGGCNAGGTCTTAGACCTATTCTCTG	971
Db	3510	AGAAATGGAATTCAGTCAATTAATGGATATGCTCCAAAGGCNAGGTCTTAGACCTATTCTCTG	3451
Qy	972	AAGCAGAAATTTGATTTGGCGGTGATTTTCATGACTACAAAGAAATTTACTGTGATCTCAGG	1031
Db	3450	AAGCAGAAATTTGATTTGGCGGTGATTTTCATGACTACAAAGAAATTTACTGTGATCTCAGG	3391
Qy	1032	GCCATCCAGTACAGGATTAAGACAAACACTCCAGGACCAAGCCTTTTCCAAAGCGGTG	1091
Db	3390	GCCATCCAGTACAGGATTAAGACAAACACTCCAGGACCAAGCCTTTTCCAAAGCGGTG	3331
Qy	1092	CAGTTGATGAAAACTTAATGCCAGCGCCCAAGTGAACACTACAACTACGTAGCTGACA	1151
Db	3330	CAGTTGATGAAAACTTAATGCCAGCGCCCAAGTGAACACTACAACTACGTAGCTGACA	3271
Qy	1152	CAGAATCAGAGCAAGCAGATACATGGG-----ATTGAGTGAAAGGCCAAAAAGAAATCA	1205
Db	3270	CAGAATCAGAGCAAGCAGATACATGGGATTTGATTTGAGTGAAGGCCAAAAAGAAATCA	3211
Qy	1206	AGTCTCCAAATGGAACAAAATTTCCAGATGCTTTTCCAGACGCCACCTCTGTAAGG	1265
Db	3210	AGTCTCCAAATGGAACAAAATTTCCAGATGCTTTTCCAGACGCCACCTCTGTAAGG	3151
Qy	1266	AGTCTCCAAACAAAGCTCTTAATAATAATAGTATGATGATCAAACTACTTTGGCTAAGATGA	1325
Db	3150	AGTCTCCAAACAAAGCTCTTAATAATAATAGTATGATGATCAAACTACTTTGGCTAAGATGA	3091
Qy	1326	GAATCCCAAACTATCAGCTTTTCCAACTAAATGCCAAGTATAAATAAAGTAAAGATA	1385
Db	3090	GAATCCCAAACTATCAGCTTTTCCAACTAAATGCCAAGTATAAATAAAGTAAAGATA	3031
Qy	1386	GGGCTTCTCAGCAGCAGCAGCAACTCCATCAGAAATCTCTTTTCCAGCGCTCTACCAAA	1445
Db	3030	GGGCTTCTCAGCAGCAGCAGCAACTCCATCAGAAATCTCTTTTCCAGCGCTCTACCAAA	2971
Qy	1446	AAAGGGAAGGGATGAAGAAATCAAGAAATGCTTTTCTCATGCAAAATCAGCAAGATAGAAA	1505
Db	2970	AAAGGGAAGGGATGAAGAAATCAAGAAATGCTTTTCTCATGCAAAATCAGCAAGATAGAAA	2911
Qy	1506	CGTCTTGTCTCTTTTAGAAACAAAACAACTGCTGCTACACCTCATTTGCGAAAAATAAGG	1565
Db	2910	CGTCTTGTCTCTTTTAGAAACAAAACAACTGCTGCTACACCTCATTTGCGAAAAATAAGG	2851
Qy	1566	AGCAGCTCTATCTGAGATGAGCCTGTGGACACAACTCAGCAAACTCAGCAAACTAATTTTACAG	1625
Db	2850	AGCAGCTCTATCTGAGATGAGCCTGTGGACACAACTCAGCAAACTCAGCAAACTAATTTTACAG	2791
Qy	1626	ATACAGATTTAAATCTATTGTGAAAAATTTCTGCCAGTAAATCTCATGCTGCGAAAAAGC	1685
Db	2790	ATACAGATTTAAATCTATTGTGAAAAATTTCTGCCAGTAAATCTCATGCTGCGAAAAAGC	2731
Qy	1686	TAAGATCAAAATAAAAAAGGGAAATTTGGATGATGTGCCCATAGAAATGAAGTATTGGAAAC	1745
Db	2730	TAAGATCAAAATAAAAAAGGGAAATTTGGATGATGTGCCCATAGAAATGAAGTATTGGAAAC	2671
Qy	1746	AGTTAATCAGGACACAAACCCAGAGTTAGAAATTTGATGTGAAAGTTTCAAAACAGGAGG	1805
Db	2670	AGTTAATCAGGACACAAACCCAGAGTTAGAAATTTGATGTGAAAGTTTCAAAACAGGAGG	2611
Qy	1806	AAAGTGTCAATGTTAGAAAAAGGCCCAAGGATGGATATAGAAAACAAATGACATTTTCAGTG	1865
Db	2610	AAAGTGTCAATGTTAGAAAAAGGCCCAAGGATGGATATAGAAAACAAATGACATTTTCAGTG	2551
Qy	1866	ATGAAGCAGTACACAGAAAGTAGCAAAATATCTCAAGAAATGAAATTTGGGAAGAACCTG	1925



Db 2550 ATGAGCAGTACCGAAGTAGCAAAATATCTCAAGAAATGAAATGGGAAGAAAGCTG 2491  
Qy 1926 AACTCAAGGAAGACTCACTATGTGCTAGCTAAAGAAATATCTAACAATGACAAACTTCAGG 1985  
Db 2490 AACTCAAGGAAGACTCACTATGTGCTAGCTAAAGAAATATCTAACAATGACAAACTTCAGG 2431  
Qy 1986 ATGATAGTAGAGTCTTCCAAAAAGAGCTGTTATTTAGCTGAATTTAGATCACTGGTGATTA 2045  
Db 2430 ATGATAGTAGAGTCTTCCAAAAAGAGCTGTTATTTAGCTGAATTTAGATCACTGGTGATTA 2371  
Qy 2046 AAAAAGCTCTCCAGAAATCCGCTGGCATAAATGATGATGATGTTGCTCAACTAAAAAAT 2105  
Db 2370 AAAAAGCTCTCCAGAAATCCATCTGGCATAAATGATGATGATGTTGCTCAACTAAAAAAT 2311  
Qy 2106 TCAAGAAATCAAAAAGGTCACATATCTCTGGAGCAGGAAAACTTCCACACATCAATGGAG 2165  
Db 2310 TCAAGAAATCAAAAAGGTCACATATCTCTGGAGCAGGAAAACTTCCACACATCAATGGAG 2251  
Qy 2166 GATCAGATCTAATAGCTCATCTGCTGAAAGAAATACAGAACTAGAAAGAGTGCGTAAGGC 2225  
Db 2250 GATCAGATCTAATAGCTCATCTGCTGAAAGAAATACAGAACTAGAAAGAGTGCGTAAGGC 2191  
Qy 2226 AGGAATGGAGGTACAAAATCAACATGCAAAAGAGAGTCTCTGCTGATGATCTTTTTA 2285  
Db 2190 AGGAATGGAGGTACAAAATCAACATGCAAAAGAGAGTCTCTGCTGATGATCTTTTTA 2131  
Qy 2286 GATACAAATCTTATTTAAAAAGGAGAGATAAAGTGGAGTATTTAAAAAGAGCCATGGAA 2345  
Db 2130 GATACAAATCTTATTTAAAAAGGAGAGATAAAGTGGAGTATTTAAAAAGAGCCATGGAA 2071  
Qy 2346 AAATCTCTAGTAGAGATCTATTTCAGGCCAAAGGTTATATGAATATATAGTGATAG 2405  
Db 2070 AAATCTCTAGTAGAGATCTATTTCAGGCCAAAGGTTATATGAATATATAGTGATAG 2011  
Qy 2406 AAGCGATTTAAGTTACAAATGTTTATGGCCCTAAATTTATTAATTAATGCACAAACTT 2465  
Db 2010 AAGCGATTTAAGTTACAAATGTTTATGGCCCTAAATTTATTAATTAATGCACAAACTT 1951  
Qy 2466 TGATCTCTTTGTATGTAACAAATGTTTGTGCTGTTTCAGGCTTTGTCATGATCTTTT 2525  
Db 1950 TGATCTCTTTGTATGTAACAAATGTTTGTGCTGTTTCAGGCTTTGTCATGATCTTTT 1891  
Qy 2526 TTTCAATTTTAAATGTTTTGTTTATTAATAGTTTAAATAGTCAAGTTCAAAATCT 2585  
Db 1890 TTTCAATTTTAAATGTTTTGTTTATTAATAGTTTAAATAGTCAAGTTCAAAATCT 1831  
Qy 2586 AAATRTACGTAAAGTAAAGACTAAAGTCAACCTTCCACCATTGCTCTAGCTACT - - - - 2640  
Db 1830 AAATATACGTAAAGTAAA - GACTAAAGTCAACCTTCCACCATTGCTCTAGCTACTTGGTT 1772  
Qy 2641 - - - - - 2640  
Db 1771 CCCTCAGAAAAAATTCATGGATCTCATTTCTTATGAATCTTTCAGGGATTTTGTAG 1712  
Qy 2641 - - - - - TATTTTAAATAATTTCTTACAAATGATAGCAATATGCA 2685  
Db 1711 TCCTATTCAAATTCCTATTATTTAAATAATTTCTTACAAATGATAGCAATATGCA 1652  
Qy 2686 GTGTTCTACACCTTGCTTTTACTTAGTAAAGATTAATAATATAGGAATATCAATATA 2745  
Db 1651 GTGTTCTACACCTTGCTTTTACTTAGT - AGAATTAATAATATAGGAATATCAATATA 1593  
Qy 2746 TGTTTTTAATATTTTCTTTTCCATTTATCTGCTAGTC - TTACTAAACTCTGCTGATCC 2804  
Db 1592 TGTTTTTAATATTTTCTTTTCCATTTATGCTGTACTTTTACCTAAACTCTGCTGATCC 1533  
Qy 2805 AAAAATAATGGCTTCAGTGGTGAGATGTCACCTACATGTTATTTCTAGTACTAGAACTG 2864  
Db 1532 AAAAATAATGGCTTCAGTGGTGAGATGTCACCTACATGTTATTTCTAGTACTAGAACTG 1473  
Qy 2865 AAGACATGTGGAGACTTCATCAAAATGCGGTTTGTGTTTTCACAGAAATGGAAAGACCTG 2924  
Db 1472 AAGACATGTGGAGACTTCATCAAAATGCGGTTTGTGTTTTCACAGAAATGGAAAGACCTG 1413

Qy 2925 TACCCCTTTTGGTGGTCTTACTAGCTGGTGGGTGCTGTGTTTGGAGCTTATTTAGAGT 2984  
Db 1412 TACCCCTTTTGGTGGTCTTACTAGCTGGTGGGTGCTGTGTTTGGAGCTTATTTAGAGT 1353  
Qy 2985 CCTAGTTTTTCTACTATTATAAGTAGAAATGGTGAGATTTGTTTCTTTTCTACCKTAAAG 3044  
Db 1352 CCTAGTTTTTCTACTATTATAAGTAGAAATGGTGAGATTTGTTTCTTTTCTACCTTAAAG 1293  
Qy 3045 GGAGATGGTAAGAAAAAATGAATGCTTTTTTCAAACTTTATTGACAAGTGATTTTCAAG 3104  
Db 1292 GGAGATGGTAAGAAAAAATGAATGCTTTTTTCAAACTTTATTGACAAGTGATTTTCAAG 1233  
Qy 3105 TCTGTGTTCAAAATAATATATTATGATCTGTGATCCAGCAAGAGGAGTTCCAGTCAAG 3164  
Db 1232 TCTGTGTTCAAAATAATATATTATGATACCTGTGATCCAGCAAGAGGAGTTCCAGTCAAG 1173  
Qy 3165 AGTCACATCAATCTGATTTAGTGTGTTAGAGAAATGGAACAGTGGAGGATGAGGC 3224  
Db 1172 AGTCACATCAATCTGATTTAGTGTGTTAGAGAAATGGAACAGTGGAGGATGAGGC 1113  
Qy 3225 CATATTTCCATGACTTTCCCTTTGTAACAGAGCAACAGAGGACAAAGAGCTGGCTCT 3284  
Db 1112 CATATTTCCATGACTTTCCCTTTGTAACAGAGCAACAGAGGACAAAGAGCTGGCTCT 1053  
Qy 3285 ACATCACTCTCACCTTTCCAAATCTTGTGGAAGTGCATCTACTTTGCCAGAACCAAAATTAAC 3344  
Db 1052 ACATCACTCTCACCTTTCCAAATCTTGTGGAAGTGCATCTACTTTGCCAGAACCAAAATTAAC 993  
Qy 3345 TTACTTCCAAAGTTCTGGCTGCTGAGGTGGAATCCAGCTGCAAGGGAGTTAGGGAAT 3404  
Db 992 TTACTTCCAAAGTTCTGGCTGCTGAGGTGGAATCCAGCTGCAAGGGAGTTAGGGAAT 933  
Qy 3405 GAAGGTCTTTTTTAAAAAGCTTCTCAGCTTCTAGGGAACAGAAATGGGTGAGCCAAAT 3464  
Db 932 GAAGGTCTTTTTTAAAAAGCTTCTCAGCTTCTAGGGAACAGAAATGGGTGAGCCAAAT 873  
Qy 3465 CTGCAATTTCTACTACAGGCATTGAGACCAGTTAGATTATTGAAATATTATAGAGAGTTA 3524  
Db 872 CTGCAATTTCTACTACAGGCATTGAGACCAGTTAGATTATTGAAATATTATAGAGAGTTA 813  
Qy 3525 TGAACACTTAATATGATAGTGGTATGACATTTGGATAGAACATGGGATCTTTAGAGT 3584  
Db 812 TGAACACTTAATATGATAGTGGTATGACATTTGGATAGAACATGGGATCTTTAGAGT 753  
Qy 3585 AGAATTTGACAGGCATTTAGTTGATGAAATGAGTCAATTTGAGTCTCTTAATAGCCATG 3644  
Db 752 AGAATTTGACAGGCATTTAGTTGATGAAATGAGTCAATTTGAGTCTCTTAATAGCCATG 693  
Qy 3645 TATCATAAATTAACAAGTGGTGGTGAACATATGGTCTCCATTTTACAGTTTAAAGGAAT 3704  
Db 692 TATCATAAATTAACAAGTGGTGGTGAACATATGGTCTCCATTTTACAGTTTAAAGGAAT 633  
Qy 3705 ATAAAGGACAGATTAATATTTGTTCTGTCATGCCCAAAATCCCTTTCTTAAGGAAGACTG 3764  
Db 632 ATAAAGGACAGATTAATATTTGTTCTGTCATGCCCAAAATCCCTTTCTTAAGGAAGACTG 573  
Qy 3765 CCTACTATAGCAGTTTTTATATTTGTCATTTATGAATATATGAATGAGTCTGAGGAGTTCTGG 3824  
Db 572 CCTACTATAGCAGTTTTTATATTTGTCATTTATGAATATATGAATGAGTCTGAGTCTGG 514  
Qy 3825 TACCTCTGCTTTTCAAAATATTTGGGTGTTGTCAGTATTTTCCCTTTTAAACMMTTC 3884  
Db 513 TACCTCTGCTTTTCAAAATATTTGG - - - - - TGTGTCAGTATTTTCTCTTTTAAACCATTC 458  
Qy 3885 CAAATTCGGGTGTAGTGGATGTTTCCATTTGGGTTTTTAATTTGATATATCCCTGATAGC 3944  
Db 457 CAAATTCGGGTGTAGTGTATG - - - - - TTTCAATTTGGTTTTAAATTTGTATATCCCTGATAGC 403  
Qy 3945 TATAATTTGGGTCAAGAAATTTCTTATACATTTAGATGCAAGTCTCTTGYCGGATATAC 4004  
Db 402 TATAATTTGGGTCAAGAAATTTCTTATACATTTAGATGCAAGTCTCTTGTGCGGATATAT 343







[illegible][illegible]



QY 3099 TTCAAGTCTGTGTTCAAAATATATTATCATGTACCTGTGATCCAGCAAGAGGAGTTCCA 3158  
Db 51896 TTCAAGTCTGTGTTCAAAATATATTATCATGTACCTGTGATCCAGCAAGAGGAGTTCCA 51955  
QY 3159 GTCAAGAGTCACTACAACATGATTAGTGTGTTAGAGAAATGAGAAATGGAACAGTGAAGAAAT 3218  
Db 51956 GTCAAGAGTCACTACAACATGATTAGTGTGTTAGAGAAATGAGAAATGGAACAGTGAAGAAAT 52015  
QY 3219 GGAGGCGCATATTTCGATGATCTCCCTTGTAAACAGAAAGCAACAGAAAGGACAGAGGCTG 3278  
Db 52016 GGAGGCGCATATTTCGATGATCTCCCTTGTAAACAGAAAGCAACAGAAAGGACAGAGGCTG 52075  
QY 3279 GCCTCTACATCACTCTCACTCTCCAAATCTGTGGAAGTGCACTACTCTTGCACAGAACAA 3338  
Db 52076 GCCTCTACATCACTCTCACTCTCCAAATCTGTGGAAGTGCACTACTCTTGCACAGAACAA 52135  
QY 3339 ATTAACCTTACTTCCAAAGTCTCTGCTCTCTGCGAGTGGAACTCCAGCTGCAAGGGAGTTAG 3398  
Db 52136 ATTAACCTTACTTCCAAAGTCTCTGCTCTCTGCGAGTGGAACTCCAGCTGCAAGGGAGTTAG 52195  
QY 3399 GGAAATGAAGGCTCTTTTAAAGCTTCTCAGCCTTCTCAGGAAACAGAAATTTGGGTGA 3458  
Db 52196 GGAAATGAAGGCTCTTTTAAAGCTTCTCAGCCTTCTCAGGAAACAGAAATTTGGGTGA 52255  
QY 3459 GCCAATCTCAATTTCTACTACAGGCATTCAGACCACTAGATATTGAAATATTATAGA 3518  
Db 52256 GCCAATCTCAATTTCTACTACAGGCATTCAGACCACTAGATATTGAAATATTATAGA 52315  
QY 3519 GAGTTATGAACACTTAAATTTATGATAGTGGTATGACATTTGGATAGAACATGGGATACTTT 3578  
Db 52316 GAGTTATGAACACTTAAATTTATGATAGTGGTATGACATTTGGATAGAACATGGGATACTTT 52375  
QY 3579 AGAAGTAGAATGACAGGCAATATTAGTTGATGAATGAATGAGATCATTTGAGTCTTAAATA 3638  
Db 52376 AGAAGTAGAATGACAGGCAATATTAGTTGATGAATGAATGAGTCTTAAATA 52435  
QY 3639 GCCATGTATCATAAATTTACCAAGTGAAGCTGGTGGAAACATATGTCCTCCATTTTACAGTTA 3698  
Db 52436 GCCATGTATCATAAATTTACCAAGTGAAGCTGGTGGAAACATATGTCCTCCATTTTACAGTTA 52495  
QY 3699 AGGAATATAATGACAGATTAATATTGTTTCTGTCATGCCCAATCCCTTTCTAAGA 3758  
Db 52496 AGGAATATAATGACAGATTAATATTGTTTCTGTCATGCCCAATCCCTTTCTAAGA 52555  
QY 3759 AGATCCGCCATCATATAGCAGTCTTTTATATTGTCATTTATGAAATATGATGAATGAGGAG 3818  
Db 52556 AGATCCGCCATCATATAGCAGTCTTTTATATTGTCATTTATGAAATATGATGAATGAGGAG 52614  
QY 3819 TTCTGGTACTCTGCTTTTACAAATATTGGGTTGTGTCAGTATTTTCCCTTTTAAAC 3878  
Db 52615 TTCTGGTACTCTGCTTTTACAAATATTGGGTTGTGTCAGTATTTTCCCTTTTAAAC 52670  
QY 3879 CMTTCCCAATTCGGGTGTGTAGTGGATGTTTCCATTTTGGGTTTAAATTTGTATATCCCT 3938  
Db 52671 CCATTCCAATTCGGGTGTGTAGTGGATGTTTCCATTTTGGGTTTAAATTTGTATATCCCT 52725  
QY 3939 GATAGCTATAATTTGGGTCTATAGAAATCTTTATACATTTAGATGCAAGTCTCTTGYCGG 3998  
Db 52726 GATAGCTATAATTTGGGTCTATAGAAATCTTTATACATTTAGATGCAAGTCTCTTGYCGG 52785  
QY 3999 ATATAGCTATTGAGATATTACACCTAGTCTGTGGCTTGACTGTTTCTTTATGCTCTTTTG 4058  
Db 52786 ATATAGCTATTGAGATATTACACCTAGTCTGTGGCTTGACTGTTTCTTTATGCTCTTTTG 52845  
QY 4059 ATGAATAGAGTCTTTTAAATTTTGACAGGTCCTTAAATTTATTTTCTTTTGTGATATT 4118  
Db 52846 ATGAATAGAGTCTTTTAAATTTTGACAGGTCCTTAAATTTATTTTCTTTTGTGATATT 52905  
QY 4119 TTTTCTCTCAATTTTAAACCCCAAGATTTGAGATATTCTGCTCTATATATAAACTTTATA 4178  
Db 52906 TTTTCTCTCAATTTTAAACCCCAAGATTTGAGATATTCTGCTCTATATATAAACTTTATA 52965  
QY 4179 TTTTATATTGATGATCACTCTTGAATTTGATATGTTGTGAAATTTATGATCGAGGTT 4238

Db 52966 TTTTATATTGATGATCACTCTTGAATTTGATATGTTGTGAAATTTATGATCAGGTT 53025  
QY 4239 CTTTTCCTCCCATCAACAGTATCCAGTATCTGTAACACTGTTTATTGAAGAATTTATCC 4298  
Db 53026 CTTTTCCTCCCATCAACAGTATCCAGTATCTGTAACACTGTTTATTGAAGAATTTATCC 53085  
QY 4299 TTTTCCTCATTAATTAATACCTTCGCAATTTAGTAAATAATCAATTAACCATRM 4348  
Db 53086 TTTTCCTCATTAATTAATACCTTCGCAATTTAGTAAATAATCAATTAACCATAA 53135  
RESULT 8  
AB013139 Homo sapiens gene for NBS1, complete cds. linear PRI 25-DEC-2002  
LOCUS AB013139  
DEFINITION AB013139.1 GI:3169124  
ACCESSION NBS1; Nijmegen breakage syndrome.  
VERSION Homo sapiens (human)  
KEYWORDS ORGANISM  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
FEATURES  
source  
exon  
5'UTR  
CDS  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8q21"  
/sub\_clones="RG255A7"  
4790..4848  
/number=1  
4790..4811  
Join(4812..4848,6518..6651,7850..7998,8480..8639,  
11050..11153,18083..18200,18816..19009,24866..24963,  
30519..30648,3818..34090,35682..36129,41481..41549,  
43078..43233,46007..46120,52298..52347,53761..53791)  
/codon\_start=1  
/product="NBS1"  
/protein\_id="BAA28616.1"  
/db\_xref="GI:3169125"  
/translation="MWKLLPAAGPAGGPPYRLLTGVYVGRKNCALIIENDQISRN  
HAVLTANFVNLSQTDIPVLTLDNSKYGTGVNNEKMQNGFSTRKSGDGIITGVF  
GSKPRIEVEPLVACSSCLDVGKALNOAILQGGTVNNWTEBCTHLVMVSKVTK  
TICALGRPIVKEPFTFEUKAVESKKOPPIESFYPLDEPSIGSKNVDLSGRQER  
KOIFPKGTFFILNAKQHKLSAVVFGGEARLITEENEEHNFPLAGTCTVDTGIT  
NSQTLIPQCKKWIOSIMDLQORLPPEABIGLAVIFMTTKNYCDPQGHPSGLK  
TTTPGSLQSGVDEKMLPAPNTTYYVADTESEQADTWDLSRPKEIKVSRMEQ  
FRLMSQAPTKVESCSTSNNSMVSNTLAKMRIETSCILLEQTQATPSLWKNKEQHL  
OTNSYRFPSTKKREDENSCKSARIETSCILLEQTQATPSLWKNKEQHL  
SENEPDTNSNNLFTDLSIKVNSKSHAAEKLRSNKKREDDVAIDEVLEOL  
FADTPELIDVKVQEDENVNRKPRMDIETNDTFSDEAVPESKISQENIECKR  
ELKEDLSMAKEISNNDKLQDSEMLPKLLLTLEFRSLVIRKNSRNSGINDDVGOL  
KNFKFKVTPGAKLPHIITGSDLIHAHARKNTELEWURQEMVQNQHAKESLSA  
DDLFRNPNLYKRR"



intron	4849..6517	Db		53812	CCATGGAAGAACTTCTAGTAAGCATCTACTTCAGGCCAACAAAGGTTATATGTAATATATA	53871
exon	/number=1	Qy		2398	GTGTATAGAAAGCGATTAAAGTTTACAAATGTTTATGCGCTTAAATTTTAAATTAATGCA	2457
intron	/number=2	Db		53872	GTGTATAGAAAGCGATTAAAGTTTACAAATGTTTATGCGCTTAAATTTTAAATTAATGCA	53931
exon	/number=2	Qy		2458	CAAAACCTTTGATCTTTTGTATGTAAACAAATGTTTGTCTGTTTTCAGGCTTTGTCTATG	2517
intron	/number=3	Db		53932	CAAAACCTTTGATCTTTTGTATGTAAACAAATGTTTGTCTGTTTTCAGGCTTTGTCTATG	53991
exon	/number=3	Qy		2518	CATCTTTTTCATTTTAAATGTTTGTATGTAAATGTTTGTATGTAAATGTTTGTATGTACAGTTC	2577
intron	/number=4	Db		53992	CATCTTTTTCATTTTAAATGTTTGTATGTAAATGTTTGTATGTAAATGTTTGTATGTACAGTTC	54051
exon	/number=4	Qy		2578	AAATTTCTAAATRTACGTAAAGGTAAAGGACTAAAGTCAACCTTCCACCACTGTCTCTAGCT	2637
intron	/number=5	Db		54052	AAATTTCTAAATRTACGTAAAGGTAAAGTCAACCTTCCACCACTGTCTCTAGCT	54110
exon	/number=5	Qy		2638	ACT-----	2640
intron	/number=6	Db		54111	ACTTGGTTCCTCCCTCAGAAAAAAATTCATGATACCTCATTTCTTATGAATCTTTCACAGGAT	54170
exon	/number=6	Qy		2641	-----TATTTTAAATAATTTCTCTACACAAATGATAGCATAAC	2678
intron	/number=7	Db		54171	TTTGTAGTCTTAAATTCCTTATTTTAAATAATTTCTCTACACAAATGATAGCATAAC	54230
exon	/number=7	Qy		2679	ATATGCAGTGTCTACACCTTGCTTTTACCTTTAGTAAGATTAAAAATTTATAGGAATATC	2738
intron	/number=8	Db		54231	ATATGCAGTGTCTACACCTTGCTTTTACCTTTAGTAAGATTAAAAATTTATAGGAATATC	54289
exon	/number=8	Qy		2739	AATATAATGTTTAAATATTTTCTTTTCCATTTATGCTGTAGTCTTACCTAAACTCTGG	2798
intron	/number=9	Db		54290	AATATAATGTTTAAATATTTTCTTTTCCATTTATGCTGTAGTCTTACCTAAACTCTGG	54349
exon	/number=9	Qy		2799	TGATCCAAACAAATGGCTTCAGTGTGCAGATGTCCACCTACATGTTTCTAGTACTAG	2858
intron	/number=10	Db		54350	TGATCCAAACAAATGGCTTCAGTGTGCAGATGTCCACCTACATGTTTCTAGTACTAG	54409
exon	/number=10	Qy		2859	AAACTGAAGACCATGTGGAGACTTTCATCAAAACATGGGTTTGTAGTTTTCACAGAAATGAAA	2918
intron	/number=11	Db		54410	AAACTGAAGACCATGTGGAGACTTTCATCAAAACATGGGTTTGTAGTTTTCACAGAAATGAAA	54469
exon	/number=11	Qy		2919	GACCTGTACCCCTTTTGGTGGTCTTAACTGAGCTGGGTGGGTGTCTGTTTGTAGCTTAT	2978
intron	/number=12	Db		54470	GACCTGTACCCCTTTTGGTGGTCTTAACTGAGCTGGGTGGGTGTCTGTTTGTAGCTTAT	54529
exon	/number=12	Qy		2979	TAGAGTCTAGTTTCTTACTTATAAAGTAGAAATGGTGCAGATTTGTTTCTTTTCTACC	3038
intron	/number=13	Db		54530	TAGAGTCTAGTTTCTTACTTATAAAGTAGAAATGGTGCAGATTTGTTTCTTTTCTACC	54589
exon	/number=13	Qy		3039	KTAAAGGGAGATGGTAAAGAAACAAATGAATGTCTTTTTCAAACTTTTATTTGACAAGTGATT	3098
intron	/number=14	Db		54590	TTAAAGGGAGATGGTAAAGAAACAAATGAATGTCTTTTTCAAACTTTTATTTGACAAGTGATT	54649
exon	/number=14	Qy		3099	TTCAAGTCTGTGTTCAAAAATATATTCATGTACTGTGATCCAGCAAGAGAGGAGTTCCA	3158
intron	/number=15	Db		54650	TTCAAGTCTGTGTTCAAAAATATATTCATGTACTGTGATCCAGCAAGAGAGGAGTTCCA	54709
exon	/number=15	Qy		3159	GTCAAGAGTCACCTACAACTGATTTAGTTGTTTGAAGATGAGAAATGGAAACAGTGAGGAAT	3218
intron	/number=16	Db		54710	GTCAAGAGTCACCTACAACTGATTTAGTTGTTTGAAGATGAGAAATGGAAACAGTGAGGAAT	54769
exon	/number=16	Qy		3219	GGAGGCCATATTTCCATGACTTCCCTTTGTAACAGAAAGCAACAGAGGAGGACAGAGGCTG	3278
intron	/number=17	Db		54770	GGAGGCCATATTTCCATGACTTCCCTTTGTAACAGAAAGCAACAGAGGAGGACAGAGGCTG	54829
exon	/number=17	Qy		3279	GCCTCTACATCACTCTCACCTTCCAAATCTTGTGGAAGTGCATCTACTTTGCCAGAACCAA	3338
intron	/number=18	Db		54830	GCCTCTACATCACTCTCACCTTCCAAATCTTGTGGAAGTGCATCTACTTTGCCAGAACCAA	54889
exon	/number=18	Qy		3339	ATTAACTTACTTCCAAAGTTCTGGCTGTGCGAGGTGGAACTCCAGCTGCAAGGGAGTTAG	3398

Query Match 42.3%; Score 1862.8; DB 9; Length 56500;  
Best Local Similarity 94.1%; Pred. No. 0;  
Matches 2024; Conservative 9; Mismatches 26; Indels 91; Gaps 6;  
Qy 2278 TCTTTTGTAGTACAACTCTTATTTTAAAGGAGAGATTAAGTCTGAGGATTTTAAAGAGAG 2337  
Db 53752 TCTTTGAGTACAACTCTTATTTTAAAGGAGAGATTAAGTCTGAGGATTTTAAAGAGAG 53811  
Qy 2338 CCATGGAAGAACTTCTCTAGTAAGCATCTACTTCAGGCCCAACAAAGTTATATGAATATATA 2397



Db	54890	ATTAACTTACTTCCAAAGTTCTGGCTGCTTGCAGTGGAACTCCAGCTGCAAGGAGTTAG	54949
QY	3399	GGAAATGAAGTCTTTTAAAGCTTCTCAGCCTTCTAGGGAACAGAAATGGGTGA	3458
Db	54950	GGAAATGAAGTCTTTTAAAGCTTCTCAGCCTTCTAGGGAACAGAAATGGGTGA	55009
QY	3459	GCCAAATCTGCAATTTCTACTACAGGCAATGAGACCAAGTTAGATTATTGAAATATTATAGA	3518
Db	55010	GCCAAATCTGCAATTTCTACTACAGGCAATGAGACCAAGTTAGATTATTGAAATATTATAGA	55069
QY	3519	GAGTTATGAACACTTAAATTTATGATAGTGGTATGACATTTGGATAGAACATGGGATCTTT	3578
Db	55070	GAGTTATGAACACTTAAATTTATGATAGTGGTATGACATTTGGATAGAACATGGGATCTTT	55129
QY	3579	AGAAATGAGATTTGACAGGGCATATTAGTTGATGAAATGAGATCTTTGAGTCTTTAAATA	3638
Db	55130	AGAAATGAGATTTGACAGGGCATATTAGTTGATGAAATGAGATCTTTGAGTCTTTAAATA	55189
QY	3639	GCCATGTTATCATAAATTTACCAAGTGAAGCTGGTGGAAACATATGCTCTCCATTTTACAGTTA	3698
Db	55190	GCCATGTTATCATAAATTTACCAAGTGAAGCTGGTGGAAACATATGCTCTCCATTTTACAGTTA	55249
QY	3699	AGGAATATAATGACAGATTAATTTGTTCTGTGTCATGCCACCAATCCCTTTCTAAGGA	3758
Db	55250	AGGAATATAATGACAGATTAATTTGTTCTGTGTCATGCCACCAATCCCTTTCTAAGGA	55309
QY	3759	AGACTGCCCTACTATAGCAGTCTTTTATATTTGTCAAATTTATGAAATATGATGAGGAG	3818
Db	55310	AGACTGCCCTACTATAGCAGTCTTTTATATTTGTCAAATTTATGAAATATGATGAGGAG	55368
QY	3819	TTCTGTGTACTCTGTCCTTTTACAAATTAATTTGGGTGTCAGTATTTTCCCTTTTAAAC	3878
Db	55369	TTCTGTGTACTCTGTCCTTTTACAAATTAATTTGGGTGTCAGTATTTTCCCTTTTAAAC	55424
QY	3879	CMTTCCTCAATTCGGGTGTGTAGTGGATGTTTCCATTTGGGTTTAAATTTGATATATCCCT	3938
Db	55425	CMTTCCTCAATTCGGGTGTGTAGTGGATGTTTCCATTTGGGTTTAAATTTGATATATCCCT	55479
QY	3939	GATAGCTATAATTTGGGTGTCATAGAAATTTCTTATACATTTCTAGATGCAAGTCTCTGCGG	3998
Db	55480	GATAGCTATAATTTGGGTGTCATAGAAATTTCTTATACATTTCTAGATGCAAGTCTCTGCGG	55539
QY	3999	ATATAGTATTGAGATATTACACCTAGTCTGCGTGTGCTGTTCTTTCTTTATGTCCTTTTG	4058
Db	55540	ATATAGTATTGAGATATTACACCTAGTCTGCGTGTGCTGTTCTTTCTTTATGTCCTTTTG	55599
QY	4059	ATGAATAGAAGTTTAAATTTTGAACAGGTCAAATTTATTTTCTTTTGTGTTGATATT	4118
Db	55600	ATGAATAGAAGTTTAAATTTTGAACAGGTCAAATTTATTTTCTTTTGTGTTGATATT	55659
QY	4119	TTTTCTCTCAATTTAACCCCAAGATTTGAGATTTCTGCTCTATTATATAAATTTATA	4178
Db	55660	TTTTCTCTCAATTTAACCCCAAGATTTGAGATTTCTGCTCTATTATATAAATTTATA	55719
QY	4179	TTTTTATATTGATCTACTCTTGAATGATGATGTTGTAATTTGTAATTTGATGATGATGATG	4238
Db	55720	TTTTTATATTGATCTACTCTTGAATGATGATGTTGTAATTTGTAATTTGATGATGATGATG	55779
QY	4239	CTTTTTTCCCCCATCAAGTATCCAGTCAATTTGAACACTGTTTATTTGAAAGAAATTTATCC	4298
Db	55780	CTTTTTTCCCCCATCAAGTATCCAGTCAATTTGAACACTGTTTATTTGAAAGAAATTTATCC	55839
QY	4299	TTTCTCTCAATTAATTTACCTTGCCTGCAATTTAGTAAATTAATTAACCATM	4348
Db	55840	TTTCTCTCAATTAATTTACCTTGCCTGCAATTTAGTAAATTAATTAACCATM	55889

RESULT 9  
AF069291/c 184919 bp DNA linear PRI 02-OCT-1998  
LOCUS Homo sapiens BAC clone 255A7 from 8q21 containing NBS1 gene,  
DEFINITION complete sequence.  
ACCESSION AF069291

AF069291.1	GI:3687828
HTG.	
Hom sapiens (human)	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 184919) Tauchi H., Matsuura S., Isomura M., Komatsu K. and Nakamura Y.
JOURNAL	Direct Submission Submitted (30-MAY-1998) Laboratory of Molecular Medicine, Institute of Medical Science, the University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
COMMENT	Mapping and sequencing information: This chromosome 8 clone was provided by S. Matsuura, H. Tauchi and K. Komatsu at the Dept. of Radiation Biology, Research Institute for Radiation Biology and Medicine, Hiroshima University, 1-2-3 Kasumi, Minami-ku, Hiroshima 734-8553, Japan. Sequencing was performed by H. Tauchi, M. Isomura and Y. Nakamura at the Human Genome Center, Institute of Medical Science, the University of Tokyo.
Source information:	
Clone 255A7	is isolated from the human BAC library available from Research Genetics, Inc. The library contains cloned DNA from the male fibroblast cell line 978SK.
Vector:	pheoBAC11
Selection:	chloramphenicol
Orientation of this clone	is centromeric (1) to telomeric (184919) side.
Location/Qualifiers	
1. .184919	
/organism="Homo sapiens"	
/mol_type="genomic DNA"	
/db_xref="taxon:9606"	
/chromosome="8"	
/map="8q21"	
/clone="255A7"	
/sex="male"	
/cell_line="978SK"	
/cell_types="fibroblast"	
60944. .86947	
/gene="ht41"	
/note="similar to Homo sapiens testis 4.1kb cDNA in GenBank Accession Number AF061326"	
Join(60944. 61201. 68647. 68801. 73157. .73293. 73635. .73826, 80160. .80251. 83583. .86947)	
/gene="ht41"	
/product="ht41"	
60944. .61201	
/gene="ht41"	
/number=1	
68647. .68801	
/gene="ht41"	
/number=2	
Join(68735. .68801. 73157. .73293. 73635. .73826, 80160. .80251, 83583. .84612)	
/gene="ht41"	
/note="unknown"	
/codon_start=1	
/product="ht41"	
/protein_id="AAC62231.1"	
/db_xref="GI:3687829"	
/translation="MPLVEETSLLEDSSVTPPVVLIIGNPSGICLSYMLSGYRVPVLSG EAIHPNTILNSKLEARHLSIVDQLEYLEGREGSSNPVAVLFDTLHPDADFGYD YPSVLHWKLEQHYIIPHVVLGKPGPGGAHNNENSLTITSFGSMWELPGKAFKDWSS KRSLLKDRVMPEIYRYKHVYGVGLQKQNFRENTITSVSLRYDQDDDDIQRDI DPPVFSNPFAGAIKNGKRGVDPVLI VGSGLTAADAVALCAYNSNIPVHFVRRR VTPSLIFKQLPKLYPEVHKVYHMMCTQSYSDSNLLSDYTSFPFHRVLSFKSDMKC YECISKEANLFGALPLVGNFVRFKGGALGVTRCLATKQKKHLFVERGGGGIA" 73157. .73293 /gene="ht41"	



```
exon      /number=3
73635..73826
/gene="hr41"
/number=4
80160..80251
/gene="hr41"
/number=5
83583..86947
/gene="hr41"
/number=6
complement (94512..143663)
/gene="NBS1"
/notes="gene for Nimegen breakage syndrome"
complement (join(94512..94692,96106..96155,102333..102446,
105220..105375,106904..106972,112324..112771,
114363..114635,117805..117934,123490..123587,
129444..129637,130253..130370,137300..137403,
139814..139973,140455..140603,141802..141935,
143605..143663))
/gene="NBS1"
/product="nibrin"
complement (94512..94692)
/number=16
complement (join(94662..94692,96106..96155,102333..102446,
105220..105375,106904..106972,112324..112771,
114363..114635,117805..117934,123490..123587,
129444..129637,130253..130370,137300..137403,
139814..139973,140455..140603,141802..141935,
143605..143663))
/gene="NBS1"
/notes="similar to yeast Xrs2"
/codon_start=1
/product="nibrin"
/db_xref="GI:3687830"
/translators="MKLLPAAAGPAGGEPYRLLTGVEYVGRKNCAILLNQDSISRN
HVLTFANFVNLSQDEIPVLTKDINSKYGTGVNEERKQNGFSRTLKSGDGIITGFGV
GKFRLEYPLVACSSLDVSGKLTALNQAILQLGGFTVNNWTEECITLVMVSVKVTIK
TICALICGRPIVKPEYIFBFLKAVESKQPPQIESYPPPLDPSIGSKNVDLSGROR
KQIFKGTIFILNAKOHKLLSSAVFGGGEARLITRENEEHNFLIACPTCVDTGIT
NSQTLIPDCQKWIQSIMDLQRLRPIPEAIEGLAVIPMTIKNYCDQHPSTGLK
TTTPGSLQGVSVDEKLPSPFVNTTIVADTESQADTWLDSERPKIIVKSMEQK
FRLMSQADPTVESCTSSNNMWSNTTIAKRIIPNYQLSPTKLPSINKSKORASQQ
QNSINRYQPSTKREDEENQMSCKSARIETSCSLEQTQPATPSLKNKEQHL
SENEPVDTSNNLFTDOLKSIKNSASKSHAAEKLRSNKKREMDVAIEDEVLQOL
FKDTKPELIDIVKQKQEDVNVKPRMDIETNDTFSDEAVPESSKISOENEIGKR
ELKEDSLSAKESINNDLQDSEMLPKLLLTETFRSLVINKSTSRNPSGINDDYGOL
KNPKPKVITYPGAGKLPHIIGGSDLIHARKNTELEWLKQEMEVEVQNQAKESLSA
DDLFRNYPYLKRRR"
complement (96106..96155)
/gene="NBS1"
/number=15
complement (102333..102446)
/gene="NBS1"
/number=14
complement (105220..105375)
/gene="NBS1"
/number=13
complement (106904..106972)
/gene="NBS1"
/number=12
complement (112324..112771)
/gene="NBS1"
/number=11
complement (114363..114635)
/gene="NBS1"
/number=10
complement (117805..117934)
/gene="NBS1"
/number=9
complement (123490..123587)
/gene="NBS1"
```

```
exon      /number=8
complement (129444..129637)
/gene="NBS1"
/number=7
complement (130253..130370)
/gene="NBS1"
/number=6
complement (137300..137403)
/gene="NBS1"
/number=5
complement (139814..139973)
/gene="NBS1"
/number=4
complement (140455..140603)
/gene="NBS1"
/number=3
complement (141802..141935)
/gene="NBS1"
/number=2
complement (143605..143663)
/gene="NBS1"
/number=1
160500..>180136
/gene="DECR"
join(160500..160641,176204..176406,177989..178046,
178166..178252,179989..>180136)
/gene="DECR"
/product="2,4-dienoyl-CoA reductase"
160500..160641
/gene="DECR"
/number=1
join(160573..160641,176204..176406,177989..178046,
178166..178252,179989..>180136)
/gene="DECR"
/codon_start=1
/product="2,4-dienoyl-CoA reductase"
/protein_id="AAC62233.1"
/db_xref="GI:3687831"
/translators="MKLLPARVFTLGSRLPCGLAPRRFSGYGTKILYQNTALQSKFF
SPLQAKMPPNSFQKQVAFITGGGTGLGKMTLLSSLGAQCQVIASRKMVDLKATAEQ
ISSOTGNKVHAIQCVRDPDMQNTVSELIKVAGHPNIVINNAAGNFISPTERLSNA
WKTITDIVLNGTAFVTLRIGKQLIKAQK"
176204..176406
/gene="DECR"
/number=2
177989..178046
/gene="DECR"
/number=3
178166..178252
/gene="DECR"

Query Match      42.3%; Score 1862.8; DB 9; Length 184919;
Best Local Similarity 94.1%; Pred. No. 0;
Matches 2024; Conservative 9; Mismatches 26; Indels 91; Gaps 6;

QY  2278  TCTTTTGTAGATAACAATCCCTTATTATAAAGAGAGAGATAAATCAGGATTTTAAAGAAG 2337
      |||||
Db   94701  TCTTTCAGATACAACTCCTTATTATAAAGAGAGAGATAAATCAGGATTTTAAAGAAG 94642

QY  2338  CCATGGAAAACTTCTTAGTAGCATCTACTTCAGGCCAACCAAGGTTATATGAATATATA 2397
      |||||
Db   94641  CCATGGAAAACTTCTTAGTAGCATCTACTTCAGGCCAACCAAGGTTATATGAATATATA 94582

QY  2398  GTGTATAGAAGCGATTAAAGTTTACAAATGTTTTATGCCTAAATTTATTATAAATAAATGCA 2457
      |||||
Db   94581  GTGTATAGAAGCGATTAAAGTTTACAAATGTTTTATGCCTAAATTTATTATAAATAAATGCA 94522

QY  2458  CAAACTTTGATCTTTTGTATGAACAATGTTTGTCTGTTTTCAGGCTTTGTCAATG 2517
      |||||
Db   94521  CAAACTTTGATCTTTTGTATGAACAATGTTTGTCTGTTTTCAGGCTTTGTCAATG 94462

QY  2518  CATCTTTTTCATTTTAAATGCTTTGTTTATTATAATAGTTTAAATAGTACACAGTTC 2577
      |||||
```



Ds	94461	CATCTTTTTCATTTTAAATGTGTTTGTGTTTATTAATAATAGTTAATATAGTCACAGTTC	94402
Qy	2578	AAATTCCTAAATRTACGTAAAGTAAGGACTAAAGTCACCCCTCCACCAATGTCCTAGCT	2637
Ds	94401	AAATTCCTAAATGTACGTAAAGSTAAA-GACTAAAGTCACCCCTCCACCAATGTCCTAGCT	94343
Qy	2638	ACT-----	2640
Ds	94342	ACTTGGTCCCTCAGAAAAAAATTCATGATACCTCAATTCCTTATGAATCTTTCCAGGGAT	94283
Qy	2641	-----TATTTTAAATAATTTCTACACAAATGATAGCATAC	2678
Ds	94282	TTTGTAGTCCTATTCAAATTCCTATTTTAAATAATTTCTACACAAATGATAGCATAC	94223
Qy	2679	ATATGAGGTGTCACACCTTGCTTTTACTTAGTAAGATTAATAATTTATAGGAATATC	2738
Ds	94222	ATATGAGGTGTCACACCTTGCTTTTACTTAGT-AGATTAATAATTTATAGGAATATC	94164
Qy	2739	AAATATAATGTTTTTAATATTTTCTTTTCCATATGCTGTAGTCTTACCTAACTCTCG	2798
Ds	94163	AAATATAATGTTTTTAATATTTTCTTTTCCATATGCTGTAGTCTTACCTAACTCTCG	94104
Qy	2799	TGATCCAAACAAATGGCTTCAGTGTGCAGATGTCACCTACATGTTATTTCTAGTACTAG	2858
Ds	94103	TGATCCAAACAAATGGCTTCAGTGTGCAGATGTCACCTACATGTTATTTCTAGTACTAG	94044
Qy	2859	AACTGAAGACCATGTGGAGACTTCATCAACATGGGTTTGTATTTTCCACGAATGGAAA	2918
Ds	94043	AAACTGAAGACCATGTGGAGACTTCATCAACATGGGTTTGTATTTTCCACGAATGGAAA	93984
Qy	2919	GACCTGACCCCTTTTGTGGTCTTACTGAGCTGGTGGTCTGTTTTCAGCTTAT	2978
Ds	93983	GACCTGACCCCTTTTGTGGTCTTACTGAGCTGGTGGTCTGTTTTCAGCTTAT	93924
Qy	2979	TAGAGTCCTAGTTTCTACTTATAAAGTAGAAATGGTGAGATGTTTCTTTTCTACC	3038
Ds	93923	TAGAGTCCTAGTTTCTACTTATAAAGTAGAAATGGTGAGATGTTTCTTTTCTACC	93864
Qy	3039	KTAAAGGGAGATGGTAAGAAAACATGAATGTCCTTTTTCAACTTTATTTGACAGTGAT	3098
Ds	93863	TTAAAGGGAGATGGTAAGAAAACATGAATGTCCTTTTTCAACTTTATTTGACAAGTGAT	93804
Qy	3099	TTCAAGTCGTGTTCAAAAATATATTATCATGTACTGTGATCCAGCAAGAGGAGTTCCA	3158
Ds	93803	TTCAAGTCGTGTTCAAAAATATATTATCATGTACTGTGATCCAGCAAGAGGAGTTCCA	93744
Qy	3159	GTCAAGAGTCACACACTGATTTAGTTGTTAGAGAAATGAGAAATGGAACAGTGAGGAAT	3218
Ds	93743	GTCAAGAGTCACACACTGATTTAGTTGTTAGAGAAATGAGAAATGGAACAGTGAGGAAT	93684
Qy	3219	GGAGGCCATATTTCCATGACTCCCTTTGTAACAGAAAGCAACAGAGGGACAGAGGCTG	3278
Ds	93683	GGAGGCCATATTTCCATGACTCCCTTTGTAACAGAAAGCAACAGAGGGACAGAGGCTG	93624
Qy	3279	GCCTCTACATCACTCTCACTCCCAATCTTTGGAAGTGCACTACTCTTCCAGAACCAA	3338
Ds	93623	GCCTCTACATCACTCTCACTCCCAATCTTTGGAAGTGCACTACTCTTCCAGAACCAA	93564
Qy	3339	ATTAACTTACTTCCAAAGTCTTGCTGCTGTCAGAGTGGAACTCCAGCTGCAAGGGATTAG	3398
Ds	93563	ATTAACTTACTTCCAAAGTCTTGCTGCTGTCAGAGTGGAACTCCAGCTGCAAGGGATTAG	93504
Qy	3399	GGAAATGAAGGCTTTTTTAAAGCTTCTCAGCCTTCTAGGGAAACAGAAATTTGGGTGA	3458
Ds	93503	GGAAATGAAGGCTTTTTTAAAGCTTCTCAGCCTTCTAGGGAAACAGAAATTTGGGTGA	93444
Qy	3459	GCCAACTGCAATTTCTACTACAGGATTTAGAGACCAAGTTAGATTTTGAATATTATAGA	3518
Ds	93443	GCCAACTGCAATTTCTACTACAGGATTTAGAGACCAAGTTAGATTTTGAATATTATAGA	93384
Qy	3519	GAGTTATGAACACTTAAATTTATGATAGTGGTATGACATTTGGATAGAACATGGGATCTTT	3578
Ds	93383	GAGTTATGAACACTTAAATTTATGATAGTGGTATGACATTTGGATAGAACATGGGATCTTT	93324
Qy	3579	AGAACTGAATTTGACAGGGCATATTAGTTGATGAATGSAGTCATTTGAGTCTTATAATA	3638
Ds	93323	AGAACTGAATTTGACAGGGCATATTAGTTGATGAATGSAGTCATTTGAGTCTTATAATA	93264
Qy	3639	GCCATGTATCATAAATTTACCAAGTGGTGGAAACATATGGTCTCCATTTTACAGTTA	3698
Ds	93263	GCCATGTATCATAAATTTACCAAGTGGTGGAAACATATGGTCTCCATTTTACAGTTA	93204
Qy	3699	AGGAATAATAAGGACAGATTAATATGTTTCTGTCTCATGCCACAAATCCCTTTCTAAGA	3758
Ds	93203	AGGAATAATAAGGACAGATTAATATGTTTCTGTCTCATGCCACAAATCCCTTTCTAAGA	93144
Qy	3759	AGAGTCCCTACTATAGCAGTTTTTATATTTTCTCAATTTATGAATATAATGAATGAGAG	3818
Ds	93143	AGAGTCCCTACTATAGCAGTTTTTATATTTTCTCAATTTATGAATATAATGAATGAGAG	93085
Qy	3819	TTCTGGTACCTCTCTCTTTTACAAATATTTGGGTGTTGTCCAGTATTTTCCCTTTTAAAC	3878
Ds	93084	TTCTGGTACCTCTCTCTTTTACAAATATTTGGGTGTTGTCCAGTATTTTCCCTTTTAAAC	93029
Qy	3879	CMTTCCTCAATTCGGGTGTAGTGGATGTTTCCATTTTGGGTTTTTAAATTTGTATATCCCT	3938
Ds	93028	CMTTCCTCAATTCGGGTGTAGTGGATGTTTCCATTTTGGGTTTTTAAATTTGTATATCCCT	92974
Qy	3939	GATAGCTATAATTTGGGTCTATAGAAATTTCTTTATACATTTCTAGATGCAAGTCTCTTGGCG	3998
Ds	92973	GATAGCTATAATTTGGGTCTATAGAAATTTCTTTATACATTTCTAGATGCAAGTCTCTTGGCG	92914
Qy	3999	ATATACGTATTTGAGATATTTACACCTAGTCTGTGGCTTGACCTGTTTCTTTATGCTTTTG	4058
Ds	92913	ATATACGTATTTGAGATATTTACACCTAGTCTGTGGCTTGACCTGTTTCTTTATGCTTTTG	92854
Qy	4059	ATGAATAGAAGTTTAAATTTTGAAGGTCAAAATTTATTTTCTTTTGTGTTGATAT	4118
Ds	92853	ATGAATAGAAGTTTAAATTTTGAAGGTCAAAATTTATTTTCTTTTGTGTTGATAT	92794
Qy	4119	TTTTCTCTCCAAATTTAAACCCCAAGATTTTCAGATATTTCTGCTCTATTATATAAATTTATA	4178
Ds	92793	TTTTCTCTCCAAATTTAAACCCCAAGATTTTCAGATATTTCTGCTCTATTATATAAATTTATA	92734
Qy	4179	TTTTTATATTTGTGATCTACCTTGAAATTTGATATGATGTTGTAATTTATGATCAGGTT	4238
Ds	92733	TTTTTATATTTGTGATCTACCTTGAAATTTGATATGATGTTGTAATTTATGATCAGGTT	92674
Qy	4239	CTTTTTCCTCCCATCAAGATATCCAGTCATTTGTAACACTGTTTATTTGAAGAATATCC	4298
Ds	92673	CTTTTTCCTCCCATCAAGATATCCAGTCATTTGTAACACTGTTTATTTGAAGAATATCC	92614
Qy	4299	TTTTCCTCATTAATTTACCTTGCCAAATTTAGTAAAAAATCAATTTAACCATRM	4348
Ds	92613	TTTTCCTCATTAATTTACCTTGCCAAATTTAGTAAAAAATCAATTTAACCATAA	92564
RESULT 10			
AF049895/c			
LOCUS	AF049895	331864 bp	DNA linear PRI 18-FEB-2004
DEFINITION	Homo sapiens chromosome 8 multiple clones map q21.3, complete sequence.		
ACCESSION	AF049895		
VERSION	AF049895.3	GI:42600578	
KEYWORDS	HTG		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 331864)		
AUTHORS	Varon, R., Viesinga, C., Platzer, M., Cerosaletti, K.M., Chrzanowska, K.H., Saar, K., Beckmann, G., Seemanova, E., Cooper, P.R., Nowak, N.J., Stumm, M., Weemes, C.M., Gatti, R.A., Wilson, R.K., Digweed, M., Rosenthal, A., Sperling, K., Concannon, P. and Reis, A. Nibrin, a novel DNA double-strand break repair protein, is mutated in Nijmegen breakage syndrome		
TITLE			



```

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Cell 93 (3), 467-476 (1998)
98250062
9590180
2 (bases 1 to 331864)
Platzter,M. and Vaton,R.
Direct Submission
Submitted (30-DEC-1998) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
Sequence update by submitter
3 (bases 1 to 331864)
Platzter,M.
Direct Submission
Submitted (04-NOV-2003) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
4 (bases 1 to 331864)
Lagemann,D. and Platzter,M.
Direct Submission
Submitted (18-FEB-2004) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
On Feb 18, 2004 this sequence version replaced gi:38153761.
----- Genome Center
Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/
Contact: gscj-submit@genome.imb-jena.de
----- Project Information
Center project name: HSNBS
Center clone name: SCB-296N11 to SCB-296C9
----- Summary Statistics
Sequencing vector: M13mpl8
Chemistry: Dye-primer ET, Dye-terminator Big Dye; 100% of
reads
Assembly program: Phrap; version 0.990329
Consensus quality: 326110 bases at least Q40
Consensus quality: 330377 bases at least Q30
Consensus quality: 331864 bases at least Q20
Quality coverage: 6.80x
-----
This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest.
-----
Neighboring sequence information:
This entry is part of a larger genomic contig. The start of this
sequence is directed towards the centromere. The start
(1..2000) overlaps with the end of the neighbouring Acc number
AF117829. The end (329871..331870) of this sequence overlaps with
the start of Acc number AF117830. It is overlapped by SCB-296N11,
SCB-316M22, SCB-296C9 and covers SCB-284N21, SCB-157K21, SCB-581I0,
SCB-228C20 entirely.
-----
This clones were finished using overlapping sequence from
accessions AC004083, AC123779, AC004612
-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
Genbank flat file format but are available as part
of this entry's ASN.1 file.
-----
Location/Qualifiers
1..331864
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```

---

```

/chromosome="8"
/map="q21.3"
/clone="multiple"
1..7008
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="SCB-296N11"
/clone_lib="Research Genetics human BAC"
1..51840
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="SCB-316M22"
/clone_lib="Research Genetics human BAC"
29365..106620
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="SCB-284N21"
/clone_lib="Research Genetics human BAC"
88317..165822
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="SCB-157K21"
/clone_lib="Research Genetics human BAC"
115700..248369
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="SCB-581I0"
/clone_lib="Research Genetics human BAC"
196689..275027
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="SCB-228C20"
/clone_lib="Research Genetics human BAC"
239072..331864
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="SCB-296C9"
/clone_lib="Research Genetics human BAC"
7009..7256
/clone="single stranded/single chemistry region"
11335..11553
/clone="single stranded/single chemistry region"
11554..11584
/clone="single clone coverage"
11556
/clone="single clone coverage"
11569
/clone="single stranded/single chemistry region"
11623
/clone="single clone coverage"
11624..11670
/clone="single clone coverage"
11671..12288
/clone="single stranded/single chemistry region"
15730..15878
/clone="pcr product sequence only"
17313..17600
/clone="single stranded/single chemistry region"
18409..18600
/clone="single stranded/single chemistry region"
18786..19030
/clone="single stranded/single chemistry region"
21357..21723
/clone="single clone coverage"
21478
/clone="single clone coverage"
21481
/clone="single clone coverage"
21500..21501
/clone="single clone coverage"
21506

```

FEATURES  
Source











[illegible]



Qy	2959	AAACTGAAGACCATGTGGAGACTTCATCAAAACATATGGGTTTAGTTTTCACCAAGATGGAAA	2918
Db	68281	AAACTGAAGACCATGTGGAGACTTCATCAAAACATATGGGTTTAGTTTTCACCAAGATGGAAA	68222
Qy	2919	GACCTGTACCCCTTTTGTGGTGCTTACTAGAGCTGGTGGGTGTCGTGTGTGTGAGCTTAAT	2978
Db	68221	GACCTGTACCCCTTTTGTGGTGCTTACTAGAGCTGGTGGGTGTCGTGTGTGTGAGCTTAAT	68162
Qy	2979	TAGAGTCCTAGTTTTCTACTTTATAAAGTAGAAAATGGTGAGATTGTGTTTTCTTTTCTACC	3038
Db	68161	TAGAGTCCTAGTTTTCTACTTTATAAAGTAGAAAATGGTGAGATTGTGTTTTCTTTTCTACC	68102
Qy	3039	KTAAAGGGAGATGGTAAAGAAAACAATGAATGTCTTTTTTCAAACCTTTATTGACAGATGATT	3098
Db	68101	TTAAAGGGAGATGGTAAAGAAAACAATGAATGTCTTTTTTCAAACCTTTATTGACAGATGATT	68042
Qy	3099	TTCAAGTCTGTGTTTCAAAAATATATTATCATGTACCTGTGATCAGCAAGAAAGGAGTTCCTCA	3158
Db	68041	TTCAAGTCTGTGTTTCAAAAATATATTATCATGTACCTGTGATCAGCAAGAAAGGAGTTCCTCA	67982
Qy	3159	GTCAGAAGTCACTACAACCTGATTAGTTGTTTATAGAAATGAGAAAATGGAACAGTGTAGGAAT	3218
Db	67981	GTCAGAAGTCACTACAACCTGATTAGTTGTTTATAGAAATGAGAAAATGGAACAGTGTAGGAAT	67922
Qy	3219	GGAGGCCATATTTCATAGACTTCCCTTGTAAACAGAAAGCAAAGAGGGCAAGAGGCTG	3278
Db	67921	GGAGGCCATATTTCATAGACTTCCCTTGTAAACAGAAAGCAAAGAGGGCAAGAGGCTG	67862
Qy	3279	GCCTCTACATCACTCTCACTCCCAATCTCTGGAAGTGCACTACTCTTGGCCAGAACCAA	3338
Db	67861	GCCTCTACATCACTCTCACTCCCAATCTCTGGAAGTGCACTACTCTTGGCCAGAACCAA	67802
Qy	3339	ATTAACTTACTTCCAAAGTTCTGGCTGCTTCAGAGTGGAACTCCAGCTGCAAGGGAGTTAG	3398
Db	67801	ATTAACTTACTTCCAAAGTTCTGGCTGCTTCAGAGTGGAACTCCAGCTGCAAGGGAGTTAG	67742
Qy	3399	GGAAATGAAGGTCTTTTTTAAAGCTTCTCAGCCTTCTAGGGAAACAGAAATTTGGGTGA	3458
Db	67741	GGAAATGAAGGTCTTTTTTAAAGCTTCTCAGCCTTCTAGGGAAACAGAAATTTGGGTGA	67682
Qy	3459	GCCAACTGCAATTTCTACTACAGGCATTTAGACCCAGTTAGATTATTGAAATATATAGA	3518
Db	67681	GCCAACTGCAATTTCTACTACAGGCATTTAGACCCAGTTAGATTATTGAAATATATAGA	67622
Qy	3519	GAGTTATGAACACTTAAATTTATGATAGTGGTATGACATTTGGATAGAAACATGGGATCTTT	3578
Db	67621	GAGTTATGAACACTTAAATTTATGATAGTGGTATGACATTTGGATAGAAACATGGGATCTTT	67562
Qy	3579	AGAAGTAGAATTCACAGGCATATTAGTTGATGAAATGGAGTCATTTGAGTCTVTTAATA	3638
Db	67561	AGAAGTAGAATTCACAGGCATATTAGTTGATGAAATGGAGTCATTTGAGTCTCTTAATA	67502
Qy	3639	GCCATGTATCATTAATTTACCAAGTGAAGCTGGTGGAAACATATGGTCTCATTATTACAGTTA	3698
Db	67501	GCCATGTATCATTAATTTACCAAGTGAAGCTGGTGGAAACATATGGTCTCATTATTACAGTTA	67442
Qy	3699	AGGAAATATATGACAGATTAAATATTGTTGTCGTGTCATGCCACAAATCCCTTTCTAAGGA	3758
Db	67441	AGGAAATATATGACAGATTAAATATTGTTGTCGTGTCATGCCACAAATCCCTTTCTAAGGA	67382
Qy	3759	AGACTGCCCTACTATAGCAGTTTTTTATATTTCTCAATTTATGAAATATTAATGAATGAGGAG	3818
Db	67381	AGACTGCCCTACTATAGCAGTTTTTTATATTTCTCAATTTATGAAATATTAATGAATGAGGAG	67323
Qy	3819	TTCTGGTACCTCTGCTCTTTAACAATATTGGGTGTGTCAGATATTTTTCCCTTTTAAAC	3878
Db	67322	TTCTGGTACCTCTGCTCTTTAACAATATTGGGTGTGTCAGATATTTTTCCCTTTTAAAC	67267
Qy	3879	CMTTTCCCAATTCGGGTGCTGAGTGGATGTTTCCATTTTGGGTTTTTAAATTTGTATATCCCT	3938
Db	67266	CCATTCCAATTCGGGTGCTGAGTGGATGTTTCCATTTTGGGTTTTTAAATTTGTATATCCCT	67212
Qy	3939	GATAGCTATAAATTTGGGCTCATAGAAAATCTTTTATACATTTAGATGCAAGTCTCTTGYCGG	3998



Db 67 CGCGGTTGCAGTCGCGCCAGCCCTGAGGAGCCGACCGATGTGGAAACTGCTGCCCGC 126

Qy 73 CGCGGCGCGGAGGAGGAGAACCATACAGACTTTTGACTGGCTGGTGAAGTACGTTGTTGG 132

Db 127 CGCGGCGCGGAGGAGGAGAACCATACAGACTTTTGACTGGCTGGTGAAGTACGTTGTTGG 186

Qy 133 AAGGAAAACTGTGCGCATTTCTAATTCGAAATGATCAGTCGATCAGCCGGAATCATGCTGT 192

Db 187 AAGGAAAACTGTGCGCATTTCTAATTCGAAATGATCAGTCGATCAGCCGGAATCATGCTGT 246

Qy 193 GTTAACTGCTAACTTTTCTGTAAACCAACCTGAGTCAAAACAGATGAAATCCCTGTATTCAG 252

Db 247 GTTAACTGCTAACTTTTCTGTAAACCAACCTGAGTCAAAACAGATGAAATCCCTGTATTCAG 306

Qy 253 ATTAAAAAGATAATCTTAAGTATGGTACCTTTGTTAATGAGGAAAAAATGCGAATGGCTT 312

Db 307 ATTAAAAAGATAATCTTAAGTATGGTACCTTTGTTAATGAGGAAAAAATGCGAATGGCTT 366

Qy 313 TTCCCGAACTTTGAAGTCCGGGGATGGTATTACTTTTGGAGTGTGTTGGAAGTAAATTCAG 372

Db 367 TTCCCGAACTTTGAAGTCCGGGGATGGTATTACTTTTGGAGTGTGTTGGAAGTAAATTCAG 426

Qy 373 AATPAGATGAGCCCTTTGGTGTGCATGCTCTCTTGTGTTAGATGCTCTGGGAAAACTGC 432

Db 427 AATPAGATGAGCCCTTTGGTGTGCATGCTCTCTTGTGTTAGATGCTCTGGGAAAACTGC 486

Qy 433 TTTAAATCAAGCTATATTCGAACTTGAGGATTTTACTGTAAACAATTTGGACGAGAATG 492

Db 487 TTTAAATCAAGCTATATTCGAACTTGAGGATTTTACTGTAAACAATTTGGACGAGAATG 546

Qy 493 CACTCACTTGTGCATGCTATCAGTGAAGTTACCATTTAAAAACAATATGTGCATCAATTTG 552

Db 547 CACTCACTTGTGCATGCTATCAGTGAAGTTACCATTTAAAAACAATATGTGCATCAATTTG 606

Qy 553 TGGAGCTCCAAATGTTAAAGCAGAAATATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAA 612

Db 607 TGGAGCTCCAAATGTTAAAGCAGAAATATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAA 666

Qy 613 GAAGCAGCTCCAAATGAAAGTTTATACCACCTCTTGATGAACCATCTATTGGAAG 672

Db 667 GAAGCAGCTCCAAATGAAAGTTTATACCACCTCTTGATGAACCATCTATTGGAAG 726

Qy 673 TAAAAATGTTGATCTGTGAGGCGGAGGAAAGAAACAAATCTTCAAAGGGGAAACATT 732

Db 727 TAAAAATGTTGATCTGTGAGGCGGAGGAAAGAAACAAATCTTCAAAGGGGAAACATT 786

Qy 733 TATATTTTGAATGCCAAACAGCATAAAGAAATGAGTTCGCGAGTTGTCTTTGGAGGTGG 792

Db 787 TATATTTTGAATGCCAAACAGCATAAAGAAATGAGTTCGCGAGTTGTCTTTGGAGGTGG 846

Qy 793 GGAAGCTAGTTGATTAACAGAGAGAAATGAAGAAACATATTTCTTTTGGCTCCGGG 852

Db 847 GGAAGCTAGTTGATTAACAGAGAGAAATGAAGAAACATATTTCTTTTGGCTCCGGG 906

Qy 853 AACGTGTTGTTGATACAGGAATAACAACTCACAGACCTTAATTCCTGACTGTCAGAA 912

Db 907 AACGTGTTGTTGATACAGGAATAACAACTCACAGACCTTAATTCCTGACTGTCAGAA 966

Qy 913 GAAATGGATTCAGTCAATATGATATGCTCCAAAGGCAAGGTCTTGAACCTATTCTCTGA 972

Db 967 GAAATGGATTCAGTCAATATGATATGCTCCAAAGGCAAGGTCTTGAACCTATTCTCTGA 1026

Qy 973 AGCAGAAATTTGAAATTCGGCGTGAATTTTTCATGACTACAAAGAAATTTACTGTGATCCTCAGGG 1032

Db 1027 AGCAGAAATTTGAAATTCGGCGTGAATTTTTCATGACTACAAAGAAATTTACTGTGATCCTCAGGG 1086

Qy 1033 CCATCCAGTACAGGATTTAAAGCAACAACTCCAGGACCAAGCTTTTCAAAGGCGGTGTC 1092

Db 1087 CCATCCAGTACAGGATTTAAAGCAACAACTCCAGGACCAAGCTTTTCAAAGGCGGTGTC 1146

Qy 1093 AGTTGATGAAAACTAATATGCAAGCGCCCGAGTGAACTACTACAAACATACGTAAGTACAC 1152

Db 1147 AGTTGATGAAAACTAATATGCAAGCGCCCGAGTGAACTACTACAAACATACGTAAGTACAC 1206

Qy 1153 AGAATCAGAGCAAGCAGATACATGCGGATTTGAGTGAAGGCCAAAGAAATCAAAGTCTC 1212

Db 1207 AGAATCAGAGCAAGCAGATACATGCGGATTTGAGTGAAGGCCAAAGAAATCAAAGTCTC 1266

Qy 1213 CAAAATGAAACAAAAATTCAGAAATGCTTTTCAAGACGACCCCACTGTAAAGGAGTCTCTG 1272

Db 1267 CAAAATGAAACAAAA----- 1282

Qy 1273 CAAAACAAGCTCTAATAATAATAGTATGGTATCAAATATCTTTGGCTAAGATGAGAAATCCC 1332

Db 1283 ----- 1282

Qy 1333 AAACTATCAGCTTTTCAACCAACTAAATTCGCAAGTATAATAAAAACTAAAGATAGGCTTTC 1392

Db 1283 -----CACCAACTAAATTCGCAAGTATAATAAAAACTAAAGATAGGCTTTC 1328

Qy 1393 TCAGAGCAGCAGACGACAACTCCATCAGAAAATCTACTTTTCAGCCGCTTACC-AAAAAAGGG 1451

Db 1329 TCAGAGCAGCAGACGACAACTCCATCAGAAAATCTACTTTTCAGCCGCTTACCANAAGGG 1388

Qy 1452 AAAGGATGAAGAAAAATCAAGAAATGTCTTATGCAAAATCAGCAAGAAATAGAAAAGTCTT 1511

Db 1389 AAAGGATGAAGAAAAATCAAGAAATGTCTTATGCAAAATCAGCAAGAAATAGAAAAGTCTT 1448

Qy 1512 GTTCTCTTTTGAACAAACACAACTGCTTACACCTCATTTGCGAAAATTAAGGAGCAGC 1571

Db 1449 GTTCTCTTTTGAACAAACACAACTGCTTACACCTCATTTGCGAAAATTAAGGAGCAGC 1508

Qy 1572 ATCTATCTGAGAAATGAGCTGTGGACACAACTCAGACAAATACTTTTACAGATACAG 1631

Db 1509 ATCTATCTGAGAAATGAGCTGTGGACACAACTCAGACAAATACTTTTACAGATACAG 1568

Qy 1632 ATTTAAAAATCTATTGTGAAAAATTTCTGCCAGTAAATCTCATGCTCGAGAAAAGTAAAGAT 1691

Db 1569 ATTTAAAAATCTATTGTGAAAAATTTCTGCCAGTAAATCTCATGCTCGAGAAAAGTAAAGAT 1628

Qy 1692 CAAATAAAAAAGGGAATGGATGATGTGGCCATAGAGATGAAGTATTGGAACAGTTAT 1751

Db 1629 CAAATAAAAAAGGGAATGGATGATGTGGCCATAGAGATGAAGTATTGGAACAGTTAT 1688

Qy 1752 TCAGGACACAAAAACAGAGTTAGAAAATTTGATGTGAAAATTTCAAAGGAGGAGGATG 1811

Db 1689 TCAAGGACACAAAAACAGAGTTAGAAAATTTGATGTGAAAATTTCAAAGGAGGAGGATG 1748

Qy 1812 TCAATGTTAGAAAAAGGCCAAGGATGATATAGAAAACAAATGACACTTTTCAGTGAAG 1871

Db 1749 TCAATGTTAGAAAAAGGCCAAGGATGATATAGAAAACAAATGACACTTTTCAGTGAAG 1808

Qy 1872 CAGTACCAAGAGTAGCAAAATATCTCAAGAAAATGAAATTCGGAAGAAAACGTGAACCTCA 1931

Db 1809 CAGTACCAAGAGTAGCAAAATATCTCAAGAAAATGAAATTCGGAAGAAAACGTGAACCTCA 1868

Qy 1932 AGGAAGACTCCTATGTCAGCTAAAGAAATATCTTAACAATGACAACTTCAGGATGATA 1991

Db 1869 AGGAAGACTCCTATGTCAGCTAAAGAAATATCTTAACAATGCGCAACTTCAGGATGATA 1928

Qy 1992 GTGAGATGCTTCCAAAAAGCTGTTATTCGACTGAATTTAGATCACTGCTGATTAAGAACT 2051

Db 1929 GTGAGATGCTTCCAAAAAGCTGTTATTCGACTGAATTTAGATCACTGCTGATTAAGAACT 1988

Qy 2052 CTACTTCCAGAAATCCGCTCGCATAAATGATGATTTATGTTCACTTAAATAATTTTC 2107

Db 1989 CTACTTCCAGAAATCCGCTCGCATAAATGATGATTTATGTTCACTTAAATAATTTTC 2044

RESULT 13  
AX875862  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS

AX875862 Sequence 10767 from Patent EP1074617.  
AX875862  
AX875862.1 GI:40030598

AX875862 2044 bp DNA linear PAT 17-DEC-2003







QY	1692	CAAAATAAAAAAGCGAAATGGATGATGTGGCCATAGAAAGTGAAGTATTGGAAACAGTTAT	1751
Db	1629	CAAAATAAAAAAGCGAAATGGATGATGTGGCCATAGAAAGTGAAGTATTGGAAACAGTTAT	1688
QY	1752	TCAAGGACACAAACACAGAGTTAGAAATTCATGATGTGAAAGTTCAAAAACAGGAGGAAGATG	1811
Db	1689	TCAAGGACACAAACACAGAGTTAGAAATTCATGATGTGAAAGTTCAAAAACAGGAGGAAGATG	1748
QY	1812	TCAATGTTGAAAAAGGCCAAGGATGATATGAACAAATGACACTTTTCAGTGATGAAG	1871
Db	1749	TCAATGTTGAAAAAGGCCAAGGATGATATGAACAAATGACACTTTTCAGTGATGAAG	1808
QY	1872	CAGTACCAGAAAGTACGCAAAATATCTCAAGAAATGAAATTTGGGAAGAAACGTGAACCTCA	1931
Db	1809	CAGTACCAGAAAGTACGCAAAATATCTCAAGAAATGAAATTTGGGAAGAAACGTGAACCTCA	1868
QY	1932	AGGAAGACTCAGTATGTCAGCTGCTGAGTAAAGAAATATCTAACATGACAAACCTTCAGGATGATA	1991
Db	1869	AGGAAGACTCAGTATGTCAGCTGCTGAGTAAAGAAATATCTAACATGACAAACCTTCAGGATGATA	1928
QY	1992	GTGAGATGCTTCCAAAAAGCTGTTATGCTGAAATTTAGATCAGCTGGTATTAATAAACT	2051
Db	1929	GTGAGATGCTTCCAAAAAGCTGTTATGCTGAAATTTAGATCAGCTGGTATTAATAAACT	1988
QY	2052	CTACTTCCAGAAATCGTCTGGCATAAATGATGATTATGCTCAACTAAAAAATTC	2107
Db	1989	CTACTTCCAGAAATCGTCTGGCATAAATGATGATTATGCTCAACTAAAAAATTC	2044
RESULT 14			
AK001017			
LOCUS	AK001017	2044 bp mRNA linear	PRI 30-JAN-2004
DEFINITION		Homo sapiens cDNA FLJ10155 fis, clone HEMBAL003433, highly similar to Homo sapiens gene for NBS1.	
ACCESSION	AK001017		
VERSION	AK001017.1	GI:7022031	
KEYWORDS		oligo capping; fis (full insert sequence).	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS		Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R., Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H., Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y., Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M., Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T., Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H., Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M., Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S., Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A., Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,P., Hara,R., Takeuchi,K., Arita,M., Inose,N., Musashino,K., Yuuki,H., Oshima,A., Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T., Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S., Teraohima,Y., Suzuki,M., Nakagawa,S., Senoh,A., Mizoguchi,H., Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T., Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M., Tashiro,H., Tanigami,A., Fujiiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T., Noguchi,S., Itoh,T., Shigetani,K., Senba,T., Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T., Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J., Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K., Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R., Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.	
TITLE		Complete sequencing and characterization of 21,243 full-length human cDNAs	

JOURNAL	Nat. Genet. 36 (1), 40-45 (2004)		
PUBMED	14702039		
REFERENCE	2		
AUTHORS	Isogai,T., Ota,T., Hayaashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sato,H., Sugano,S., Shiratori,A., Sudo,H., Wagatsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K., Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahara,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.		
TITLE	NEDO human cDNA sequencing project		
JOURNAL	Unpublished		
REFERENCE	3 (bases 1 to 2044)		
AUTHORS	Isogai,T. and Otsuki,T.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)		
COMMENT	NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.		
FEATURES	Location/Qualifiers		
source	1..2044		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="HEMBAL003433"		
	/tissue_type="whole embryo, mainly head"		
	/clone_lib="HEMBAL"		
	/dev_stage="embryo, 10 weeks"		
misc_feature	/note="cloning vector: pME18SFL3"		
	1282..1283		
	/note="118 bases segment is present in AF051334 and AF058696"		
misc_difference	1378		
	/note="compared to AF051334 and AF058696"		
	/replace=""		
ORIGIN			
Query Match	41.7%; Score 1836.4; DB 9; Length 2044;		
Best Local Similarity	94.3%; Pred. No. 0;		
Matches 1976; Conservative	0; Mismatches 1; Indels 119; Gaps 2;		
QY	13	CGCGGTTGCACGTCCGCCCCAGCCCTGAGGAGCCGGACCGATGTGGAAACTGCTGCCCGC	72
Db	67	CGCGGTTGCACGTCCGCCCCAGCCCTGAGGAGCCGGACCGATGTGGAAACTGCTGCCCGC	126
QY	73	CGCGGCGCGGCGAGGAGAACCATACAGACTTTTGTGACTGGCGTTGAGTAGCTGTGG	132
Db	127	CGCGGCGCGGCGAGGAGAACCATACAGACTTTTGTGACTGGCGTTGAGTAGCTGTGG	186
QY	133	AAGGAAAACTGTGCCATTCTTAATGAAAATGATCAGTCGATCAGCCGAAATCATGCTGT	192
Db	187	AAGGAAAACTGTGCCATTCTTAATGAAAATGATCAGTCGATCAGCCGAAATCATGCTGT	246
QY	193	GTTAACTGCTAACTTTTCTGTAACCAACTGAGTCAAAACAGATGAAATCCCTGTATTGAC	252
Db	247	GTTAACTGCTAACTTTTCTGTAACCAACTGAGTCAAAACAGATGAAATCCCTGTATTGAC	306
QY	253	ATTAAGAATAATTCATGATGTTACCTTTGTTTAAATGAGGAAAAATGCGAATGGCTT	312
Db	307	ATTAAGAATAATTCATGATGTTACCTTTGTTTAAATGAGGAAAAATGCGAATGGCTT	366
QY	313	TTCCCGAACTTTGAAGTCGGGGATGGTATTACTTTTCGAGTGTTTGGAAGTAAATTCAG	372
Db	367	TTCCCGAACTTTGAAGTCGGGGATGGTATTACTTTTCGAGTGTTTGGAAGTAAATTCAG	426
QY	373	AATAGATATGAGCCCTTGGTTGTCATGCTCTTCTTTGTATTAGATGTCCTCTGGGAAACTGC	432







CDS	/gene="NBS1" 40..2292 /function="required for MRE11, RAD50 and NBS1 colocalization at nuclear foci in response to DNA double-strand breaks" /note="contains forkhead-associated (FHA) and breast cancer carboxy terminal (BRCT) domains; mutated in Nijmegen breakage syndrome" /codon_start=1 /product="Nbs1" /protein_id="AAF91228.1" /db_xref="GI:9651648"
	/translation="NM_001171AAAPGPRCLLAGVEYIVGRKNCALIIENDQISIRN HAVLRVNPVTLSOTDEIPLTI IKONGKYTFINEKRMGLSLTKLTDGRVTFGVF ESKFRVEPLVSCSLDVSCKTVLNQAILQLGLTANSWTERBCHNLMSVKVITIK TICALLCGRIPIKVCFYSEFLKAVESKTQPPEIESFPYPIDEPALGNKSVLDSSRRR KQIFKRTVFVNAKQHKLKGSAVFGGEARKLMABGEESFFSPAGTCVVDVGIT NTQILIDISQRWIHLIMDILORHDLRPITPAEIGLAVIFMTYESYNPQQPCTEYK TTTPGPSLGSLSANGKVPISPMNMWTTPADTESPADTCWSERPPEEVKI FGLDQ QNSIKNYQPCKSRKEEDENPEOSSCKSRVSNLYQLSPMKCPAASKNKWSSOQ QSTELGRAASNVGGIDIKPKNGKSPDKSFTEDLRARKEVDLTSEBEVLLELLR STKPELVAVKEQEKADVS IRKKPRMDAERNQHLNGVPPESNSALQOEDEKKOSL QIBAWETREVSNTELQPSSELPRKLLDTFRLSVLNNNSRNLCPLANGRGKLPN KFKPKATCFGAGKLPHI IGGSDLI GHHARKNTLEBWLKHMEVQKQAQEDSLADDL FRYPNVKER"
ORIGIN	
Query Match	32.2%; Score 1416.4; DB 10; Length 2605;
Best Local Similarity	76.1%; Pred. No. 3e-243;
Matches 1863; Conservative	0; Mismatches 546; Indels 39; Gaps 8;
Qy	37 CTGAGGACCGGACCATGTGGAACTGCTGCGCGCGCGGCCCGGCAGGAGGACC 96
Dd	
Qy	24 CTGAGACGCGGACTCATGTGGAACTGCTGCGCGCGCGCAAGTCGGCTCCAGGAGACC 83
Dd	
Qy	97 ATACAGACTTTTGACTGGCGCTTGAGTACCGTGTGGAAGGAAAAAATCTGCGCATTCTAAT 156
Dd	
Qy	84 ATGTCGACTTTTGGCTGGCGTAGAGTACATTGTTGGAGGAAAAAATCTGCGCATTCTAAT 143
Dd	
Qy	157 TGAATAATGATCAGTCGATCAGCGCAAATCATGCTGTGTAACTGCTAACTTTTCGTAAAC 216
Dd	
Qy	144 CGAAAATGATCAGTCAATCAGTCGAAATCATGCTGTATTAAAGATAAACTTTTCTGTAAAC 203
Dd	
Qy	217 CAACCTGAGTCAACACAGATGAATCCCTGTATTGTGACATTAAGAATTAATCTTAAGTATGG 276
Dd	
Qy	204 CAGTTTGAGTCAACACGATGAATTCCTACATTAATAATAAAGATTAATCTTAAGTATGG 263
Dd	
Qy	277 TACCTTTGTTAATGAGAAAAAATGCAGAAATGGCTTTTCCCAGAACTTTGAAAGTCGGGGA 336
Dd	
Qy	264 AACCTTTATTAAACAAGAAAAAATGCAGAAATGGCTTTTCTTCCAATTTGAACACAGGAGA 323
Dd	
Qy	337 TGGTATTACTTTTTGGAGTGTGTTGGAAGTAAATTCAGAAATAGAGTATGAGCCCTTTGGTTGC 396
Dd	
Qy	324 TAGAGTTTACCTTTGGGGTGTGTTGAAAGTAAATTCAGAGTAGAATACGAGCCCTTTGGTTGT 383
Dd	
Qy	397 ATGCTCTCTGTTTATAGTCTCTGGGAAAAACTGCTTTTAAATCAAGCTATATTGCAACT 456
Dd	
Qy	384 TTGCTCTCTGTTTATAGTCTCTGGGAAAAACTGTTTAAACCAAGCTATATTACAGCT 443
Dd	
Qy	457 TGGAGGATTTTACTGTATAAACAAATTTGGACACAGAATGCACTCACCTTTGTCAGTATTACAGT 516
Dd	
Qy	444 TGGAGGACTCAGTCGAAAACAGCTGGACAGAGAGTGTACTCACCTTGCATGTCATCAGT 503
Dd	
Qy	517 GAAAGTTTACCAATTAACCAATATGTGCATCTATTTGTGGACGTCCAATTTGTAAGCCAGA 576
Dd	
Qy	504 CAAAGTCACCAATTAACCAATATATGTGCATCTATTTGTGGACGTCCAATTTGTAAGCCAGA 563
Dd	
Qy	577 ATATTTTACTGAAATTTCTGAAAGCAGTTTCAGTCCAAAGAAGGAGCCCTCCACAAATTTGAAAG 636
Dd	
Qy	564 ATATCTTTCTGAAATTTCTCAAGACAGTTTGAAGTATTCGAAACAGTTATTCAAGGACACAAAACAGAGTT 623
Dd	







**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2005, 04:04:08 ; Search time 2151 Seconds  
(without alignments)  
12117.444 Million cell updates/sec

Title: US-09-837-138-1  
Perfect score: 4403  
Sequence: 1 ttccggcacagcgcggttg.....accgcggtgagctccagct 4403

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 4390206 segs, 2959870667 residues  
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4123.4	93.6	4412	8 ABZ72011	Abz72011 Human nib
2	4105.2	93.2	4386	3 AAZ89047	Aaz89047 Human nib
3	4103.6	93.2	4386	3 AAZ34997	Aaz34997 NBS1 gene
4	1862.8	42.3	65921	3 AAZ89046	Aaz89046 Human nib
5	1836.4	41.7	2044	4 AAH13813	Aah13813 Human cdn
6	671	15.2	752	4 AAH03660	Aah03660 Human cdn
7	564.6	12.8	622	8 ABZ71727	Abz71727 Human can
8	552.2	12.5	646	3 AAA78056	Aaa78056 cdna enco
9	552.2	12.5	646	4 AAI28794	Aai28794 Colon tum
10	552.2	12.5	646	8 ABZ32980	Abz32980 Human col
11	544.8	12.4	561	6 ABQ57768	Abq57768 Human col
12	538.6	12.2	544	6 ABL38148	Ab138148 Human col
13	520.4	11.8	543	4 AAH09188	Aah09188 Human kid
14	507	11.5	587	7 ADS71935	Ads71935 Human kid
15	405.6	9.2	468	9 ACH26622	Ach26622 Human adu
16	402.2	9.1	610	13 ADQ79117	Adq79117 Novel can
17	362	8.2	483	6 ABQ58279	Abq58279 Human col
18	358.4	8.1	396	4 AAF44923	Aaf44923 Human bre
19	239.4	5.4	272	7 ADS70151	Ads70151 Corn seed
20	212.6	4.8	475	6 ABK62896	Abk62896 Rat seque

c	21	212.6	4.8	475	12	ADP72176	Adp72176 Renal tox
	22	193	4.4	195	2	AAT24560	Aat24560 Human gen
	23	167.6	3.8	315	8	ABX43562	Abx43562 Bovine ES
	24	163.4	3.7	6440	4	AAK68763	Aak68763 Human imm
	25	163.4	3.7	6440	4	AAK68764	Aak68764 Human imm
	26	163.4	3.7	6440	4	AAK68762	Aak68762 Human imm
	27	120.2	2.7	121	2	AAT25563	Aat25563 Human gen
	28	119.4	2.7	442	4	AAI38231	Aai38231 Probe #69
	29	119.4	2.7	442	4	AAK32388	Aak32388 Human bon
	30	119.4	2.7	442	4	AAK06689	Aak06689 Human bon
	31	119.4	2.7	442	4	ABS32095	Abs32095 Human liv
	32	119.4	2.7	442	6	ABS07171	Abs07171 Human gen
	33	114	2.6	114	4	AAI51546	Aai51546 Probe #20
	34	114	2.6	114	4	AAK45612	Aak45612 Human bon
	35	114	2.6	114	4	AAK19606	Aak19606 Human bra
	36	114	2.6	114	4	ABS45301	Abs45301 Human liv
	37	114	2.6	114	6	ABS19884	Abs19884 Human gen
	38	97.6	2.2	6237	6	ABL32358	Ab132358 Human imm
	39	92.4	2.1	6237	6	ABL32359	Ab132359 Human imm
c	40	87.6	2.0	61739	12	ADQ97343	Adq97343 Human can
c	41	81.6	1.9	3285	7	ADS31392	Ads31392 Human gen
	42	80	1.8	1300	12	ADP85917	Adp85917 Synthetic
c	43	73.6	1.7	539	11	ACN80219	Acn80219 Breast ca
	44	73.6	1.7	228835	12	ADQ97421	Adq97421 Human can
	45	72.2	1.6	7442	4	AA546686	Aas46686 Tumour au

ALIGNMENTS

RESULT 1  
ABZ72011  
ID ABZ72011 standard; cdna; 4412 BP.  
XX  
AC ABZ72011;  
XX  
DT 01-APR-2003 (first entry)  
XX  
DE Human nibrin (NBS) cdna GenBank AF051334.  
XX  
KW Human; cancer; stomach cancer; cytostatic; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200283899-A1.  
XX  
PD 24-OCT-2002.  
XX  
PF 28-MAR-2002; 2002WO-JF003038.  
XX  
PR 10-APR-2001; 2001JP-00112039.  
PR 21-SEP-2001; 2001JP-00290193.  
XX  
(TAKA-) TAKARA BIO INC.  
PA Yoshikawa Y, Okamoto S, Oura T, Mineno J, Asada K, Kato I;  
PI Inoue H, Mori M;  
XX  
WPI; 2003-093022/08.

Measuring changes in expression of 264 cancer associated genes for detection of stomach cancer and screening of potential anticancer agents.  
Claim 2; Page; 266pp; Japanese.

The invention relates to a method for the detection of cancer in which a change in the expression of 1 or more of 264 specified cancer associated genes, ABZ71694-ABZ71957, or of sequences at least 80% homologous to them in the specimen tissue as compared to normal tissue is observed. The genes are used in detection, diagnosis and treatment of cancer, especially of stomach cancer. The present sequence is that of a cancer associated polynucleotide of the invention. Note: The present sequence was not given in the printed specification but was isolated using the



CC	GenBank accession number given in the DE line	
XX	Sequence 4412 BP; 1470 A; 753 C; 853 G; 1336 T; 0 U; 0 Other;	
SQ	Query Match 93.6%; Score 4123.4; DB 8; Length 4412; Best Local Similarity 97.1%; Pred. No. 0; Matches 4287; Conservative 7; Mismatches 28; Indels 91; Gaps 6;	
QY	13 CGCGGTTGACGTGCGGCCCGCCCTGAGAGCGCGACCGATGTGAAAACGTGTCGCCGC 72	Db
DB	12 CGCGGTTGACGTGCGGCCCGCCCTGAGAGCGCGACCGATGTGAAAACGTGTCGCCGC 71	QY
QY	73 CGCGGCCCGGACGAGGAGAACCATACAGACTTTTGGCTGGCTTGGTACGTTGCG 132	Db
DB	72 CGCGGCCCGGACGAGGAGAACCATACAGACTTTTGGCTGGCTTGGTACGTTGCG 131	QY
QY	133 AAGGAAAACTGTGCCATTCTTAATTGAAAATGATCGATCAGCGCGAAATCATGCTGT 192	Db
DB	132 AAGGAAAACTGTGCCATTCTTAATTGAAAATGATCGATCAGCGCGAAATCATGCTGT 191	QY
QY	193 GTTAACGTGTAATCTTTCTGTAAACCAACCTGAGTCAAAACAGATGAAATCCCTGTATTGAC 252	Db
DB	192 GTTAACGTGTAATCTTTCTGTAAACCAACCTGAGTCAAAACAGATGAAATCCCTGTATTGAC 251	QY
QY	253 ATTTAAAGATTAATCTTAAGTATGTTACCTTTTGTAAATGAGGAAAAAATGCAGAAATGGCTT 312	Db
DB	252 ATTTAAAGATTAATCTTAAGTATGTTACCTTTTGTAAATGAGGAAAAAATGCAGAAATGGCTT 311	QY
QY	313 TTCCCGAACTTTGAACTCGGGGATGGTATTACTTTTGGAGTGTTTTGGAGTAAATTCAG 372	Db
DB	312 TTCCCGAACTTTGAACTCGGGGATGGTATTACTTTTGGAGTGTTTTGGAGTAAATTCAG 371	QY
QY	373 AATAGATGATGACCTTTGGTTCGATGCTCTTCTGTTAGATGCTCTGGGAAACTGC 432	Db
DB	372 AATAGATGATGACCTTTGGTTCGATGCTCTTCTGTTAGATGCTCTGGGAAACTGC 431	QY
QY	433 TTTAAATCAAGCTATATGCACTCGAGGATTTACTGTAAACAAATTTGACAGAGAAATG 492	Db
DB	432 TTTAAATCAAGCTATATGCACTCGAGGATTTACTGTAAACAAATTTGACAGAGAAATG 491	QY
QY	493 CACTCACTTTGTCATGGTATCAGTGAAGTTACCATTAATAACAATAATGTGCACTCATTTG 552	Db
DB	492 CACTCACTTTGTCATGGTATCAGTGAAGTTACCATTAATAACAATAATGTGCACTCATTTG 551	QY
QY	553 TGGACGTCAATTTGTAAGCCAGAAATATTTTACTGAAATTCCTGAAAGCAATTCAGTCCAA 612	Db
DB	552 TGGACGTCAATTTGTAAGCCAGAAATATTTTACTGAAATTCCTGAAAGCAATTCAGTCCAA 611	QY
QY	613 GAAGCAGCTCCACAAATTTGAAAGTTTACCACCTCTTGATGAACCATCTATTGGAAG 672	Db
DB	612 GAAGCAGCTCCACAAATTTGAAAGTTTACCACCTCTTGATGAACCATCTATTGGAAG 671	QY
QY	673 TAAAAATGTTGATCTGTGAGGACGGACGGAAGAAACAAATCTTCAAAGGGAAAAACATT 732	Db
DB	672 TAAAAATGTTGATCTGTGAGGACGGACGGAAGAAACAAATCTTCAAAGGGAAAAACATT 731	QY
QY	733 TATATTTTGAATGCCAAACAGCATAGAAATTTGAGTTCGCGAGTTGCTCTTTGGAGGTGG 792	Db
DB	732 TATATTTTGAATGCCAAACAGCATAGAAATTTGAGTTCGCGAGTTGCTCTTTGGAGGTGG 791	QY
QY	793 GGAAGCTAGCTGTATACAGAGAGAAATGAAGAAGAAACATAATTTCTTTTGGCTCCGGG 852	Db
DB	792 GGAAGCTAGCTGTATACAGAGAGAAATGAAGAAGAAACATAATTTCTTTTGGCTCCGGG 851	QY
QY	853 AACGTGTGTTGTATACAGGATTAACAACTCACAGACTTAAATTCCTGACTGTGAGAA 912	Db
DB	852 AACGTGTGTTGTATACAGGATTAACAACTCACAGACTTAAATTCCTGACTGTGAGAA 911	QY
QY	913 GAAATGGATTCAGTCAATTAATGGATATGCTCCAAAGGCAAGGTCTTAGACCTTATTCCTGA 972	Db
DB	912 GAAATGGATTCAGTCAATTAATGGATATGCTCCAAAGGCAAGGTCTTAGACCTTATTCCTGA 971	QY
QY	973 AGCAGAAATTTGGATTGGCGGTGATTTTCATGACTACAAAGAAATTTACTGTGATCCTCAGGG 1032	Db
DB	972 AGCAGAAATTTGGATTGGCGGTGATTTTCATGACTACAAAGAAATTTACTGTGATCCTCAGGG 1031	QY
QY	1033 CATCCCTAGTACAGGATTAAGACAACTCCAGGACCAAGCCCTTTTCACAAGGGGTGTC 1092	Db
DB	1032 CATCCCTAGTACAGGATTAAGACAACTCCAGGACCAAGCCCTTTTCACAAGGGGTGTC 1091	QY
QY	1093 AGTTGATGAAAACCTAATGCCAAGCGCCCTGAGTGAACACTCAACATAGTGTAGTCACAC 1152	Db
DB	1092 AGTTGATGAAAACCTAATGCCAAGCGCCCTGAGTGAACACTCAACATAGTGTAGTCACAC 1151	QY
QY	1153 AGAATCAGACGACGAGATACATGGGATTTTGAAGTGAAGGCCCAAGAAATCAAAAGTCTTC 1212	Db
DB	1152 AGAATCAGACGACGAGATACATGGGATTTTGAAGTGAAGGCCCAAGAAATCAAAAGTCTTC 1211	QY
QY	1213 CAAAATGGAACAAAATTCAGAAATGCTTTTCAAGACGACCCACTGTGTAAGAGGTCTCTG 1272	Db
DB	1212 CAAAATGGAACAAAATTCAGAAATGCTTTTCAAGATGACCCACTGTGTAAGAGGTCTCTG 1271	QY
QY	1273 CAAAACAGCTCTTAATAATAATAGTATGGTATCAATACTTTGGCTTAAGATCAGAAATCCC 1332	Db
DB	1272 CAAAACAGCTCTTAATAATAATAGTATGGTATCAATACTTTGGCTTAAGATCAGAAATCCC 1331	QY
QY	1333 AAACTATCAGCTTTTCAACCAACTAAATTTGCCAAGTATAAATAAAGTAAAGATAGGGCTTC 1392	Db
DB	1332 AAACTATCAGCTTTTCAACCAACTAAATTTGCCAAGTATAAATAAAGTAAAGATAGGGCTTC 1391	QY
QY	1393 TCAGACGACGACGACCAACTCCATCAGAAAATCTTTTCAAGCGACCCACTGTGTAAGAGGTCTCTG 1452	Db
DB	1392 TCAGACGACGACGACCAACTCCATCAGAAAATCTTTTCAAGCGACCCACTGTGTAAGAGGTCTCTG 1451	QY
QY	1453 AAGGATGAAAGAAATCAAGAAATGCTTTTCAAGAAATGCTTTTCAAGCGACCAAGTAAAGTCTCTG 1512	Db
DB	1452 AAGGATGAAAGAAATCAAGAAATGCTTTTCAAGAAATGCTTTTCAAGCGACCAAGTAAAGTCTCTG 1511	QY
QY	1513 TTCTCTTTTAGAACAAACCAACCTGCTACACCTCATTTGTGGAAAAATTAAGGACGACCA 1572	Db
DB	1512 TTCTCTTTTAGAACAAACCAACCTGCTACACCTCATTTGTGGAAAAATTAAGGACGACCA 1571	QY
QY	1573 TCTATCTGAGAAATGAGCTGTGGACACAAAATCTCAGACAAATAACTTATTTTACAGATACAGA 1632	Db
DB	1572 TCTATCTGAGAAATGAGCTGTGGACACAAAATCTCAGACAAATAACTTATTTTACAGATACAGA 1631	QY
QY	1633 TTTTAAATCTATTTGTGAAAAATTTCTGCGAGTAAATCTCATGCTCAGAAAAAGCTAAGATC 1692	Db
DB	1632 TTTTAAATCTATTTGTGAAAAATTTCTGCGAGTAAATCTCATGCTCAGAAAAAGCTAAGATC 1691	QY
QY	1693 AAATAAAAAGGGAAATGGATGATGTGGCCATAGAAGATGAAGTATTGGAAACAGTTATT 1752	Db
DB	1692 AAATAAAAAGGGAAATGGATGATGTGGCCATAGAAGATGAAGTATTGGAAACAGTTATT 1751	QY
QY	1753 CAAAGACACAAAACAGAGTTAGAAATTTGATGTGAAAGTTCAAAAACAGGAGGAAGATGT 1812	Db
DB	1752 CAAAGACACAAAACAGAGTTAGAAATTTGATGTGAAAGTTCAAAAACAGGAGGAAGATGT 1811	QY
QY	1813 CAATGTTAGAAAAAGCCCAAGATGATATAGAAAACAAATGACATTTTCAAGTGAAGC 1872	Db
DB	1812 CAATGTTAGAAAAAGCCCAAGATGATATAGAAAACAAATGACATTTTCAAGTGAAGC 1871	QY
QY	1873 AGTACCAAGAAAGTAGCAAAAATCTTCAAGAAAAATGAAATTTGGGAAGAAAACGTGAACCTCAA 1932	Db
DB	1872 AGTACCAAGAAAGTAGCAAAAATCTTCAAGAAAAATGAAATTTGGGAAGAAAACGTGAACCTCAA 1931	QY
QY	1933 GGAAGACCTCATGATGGTCACTTAAAGAAATATCTTAACAAATGACAACTTCAGGATGATAG 1992	Db
DB	1932 GGAAGACCTCATGATGGTCACTTAAAGAAATATCTTAACAAATGACAACTTCAGGATGATAG 1991	QY
QY	1993 TCAGATGCTTCCAAAAAGCTGTTATTCAGTCAATTTTATAGTCACTGGTGTATTAAGAACTC 2052	Db
DB	1992 TCAGATGCTTCCAAAAAGCTGTTATTCAGTCAATTTTATAGTCACTGGTGTATTAAGAACTC 2051	QY
QY	2053 TACTTCCAGAAATCCGCTCTGGCATAAATGATGATTTATGGTCAACTTAAAAAAATTTCAAGAA 2112	Db



Db 2052 TACTTCAGAAATCCATCTCGCATAAATGATGATTTATGGTCAACTAAAAATTTCAAGAA 2111  
Qy 2113 ATTCAAAAAGGTACATATCTCGAGCAGGAAATCTCCACACATCATTTGGAGGATCAGA 2172  
Db 2112 ATTCAAAAAGGTACATATCTCGAGCAGGAAATCTCCACACATCATTTGGAGGATCAGA 2171  
Qy 2173 TCTAAATAGCTCATCATGCTCGAAAGAAATACAGAACTAGAAAGAGTGGCTAAGCGAGAAAT 2232  
Db 2172 TCTAAATAGCTCATCATGCTCGAAAGAAATACAGAACTAGAAAGAGTGGCTAAGCGAGAAAT 2231  
Qy 2233 GGAGGTACAAAAATCAACATCCAAAGAAAGAGTCTCTTGCTGATGATCTTTTATAGATACAA 2292  
Db 2232 GGAGGTACAAAAATCAACATCCAAAGAAAGAGTCTCTTGCTGATGATCTTTTATAGATACAA 2291  
Qy 2293 TCCTTATTTAAAAAGGAGAAATAACTGAGGATTTTAAAAAGAGCCATCGAAAAACTTTC 2352  
Db 2292 TCCTTATTTAAAAAGGAGAAATAACTGAGGATTTTAAAAAGAGCCATCGAAAAACTTTC 2351  
Qy 2353 CTAGTAAGCATCTACTTCAGGCCAAACAAAGGTATATAGAAATATATAGTGTATAGAGCGAT 2412  
Db 2352 CTAGTAAGCATCTACTTCAGGCCAAACAAAGGTATATAGAAATATATAGTGTATAGAGCGAT 2411  
Qy 2413 TTAAGTTACAAATGTTTATGGCCTAAATTTATTAATAAATGCACAAAATTTGATTTCT 2472  
Db 2412 TTAAGTTACAAATGTTTATGGCCTAAATTTATTAATAAATGCACAAAATTTGATTTCT 2471  
Qy 2473 TTTGATGTAAACAAATGTTTGTCTGTTTTCAGGCTTTGTCATTTGTCATCTTTTTCATT 2532  
Db 2472 TTTGATGTAAACAAATGTTTGTCTGTTTTCAGGCTTTGTCATTTGTCATCTTTTTCATT 2531  
Qy 2533 TTTTAAATGCTTTTGTATTAAATAGTTAAATATAGTCACAGTTTCAAAAATTTCTAAATRTA 2592  
Db 2532 TTTTAAATGCTTTTGTATTAAATAGTTAAATATAGTCACAGTTTCAAAAATTTCTAAATRTA 2591  
Qy 2593 CGTAAGGTAAGACATTAAGTCAACCTTCCACCAATGTCCTAGTACT----- 2640  
Db 2592 CGTAAGGTAAGACATTAAGTCAACCTTCCACCAATGTCCTAGTACTTTCCCTCCCTCA 2650  
Qy 2641 ----- 2640  
Db 2651 GAAAAAATTCATGACTACTCATTTCTTATGAATCTTCCAGGATTTTTCAGTCCCTATTTC 2710  
Qy 2641 -----TATTTTAAATATTTTCTACACAAATGATAGATACATATGCAAGTTCCTA 2693  
Db 2711 AAATTCCTATTTTAAATATTTTCTACACAAATGATAGTAAATATGCAAGTTCCTA 2770  
Qy 2694 CACCTTGCTTTTACTAGTAAATTAAGGATTAAGGATTAATCAATATAATGTTTTTA 2753  
Db 2771 CACCTTGCTTTTACTTAGT-AGATTTAAATTTATAGGAATATCAATATAATGTTTTTA 2829  
Qy 2754 ATATTTTCTTTTCCATATGCTGTAGTCTTACCTAAACTCTGGTGATCCAAACAAAAT 2813  
Db 2830 ATATTTTCTTTTCCATATGCTGTAGTCTTACCTAAACTCTGGTGATCCAAACAAAAT 2889  
Qy 2814 GGCTTCAGTGGTGCAGATGTCACCTACATGTTATTTCTAGTACTAGAACTGAAGACCATG 2873  
Db 2890 GGCTTCAGTGGTGCAGATGTCACCTACATGTTATTTCTAGTACTAGAACTGAAGACCATG 2949  
Qy 2874 TGAGACTTCATCAACATGGGTTTGTAGTTTTCACAGAAATGGAAGACCTGTACCCCTTT 2933  
Db 2950 TGGAGACTTCATCAACATGGGTTTGTAGTTTTCACAGAAATGGAAGACCTGTACCCCTTT 3009  
Qy 2934 TTGGTGTCTTACTGAGCTGGGTGCTGTTTGTAGCTTATTTAGAGTCCCTAGTTTT 2993  
Db 3010 TTGGTGTCTTACTGAGCTGGGTGCTGTTTGTAGCTTATTTAGAGTCCCTAGTTTT 3069  
Qy 2994 CCTACTTATAAAGTAGAAATGGTGAGATTTGTTTTCTTTCTACCTACCAAGGAGGATGGT 3053  
Db 3070 CCTACTTATAAAGTAGAAATGGTGAGATTTGTTTTCTTTCTACCTTAAAGGAGGATGGT 3129  
Qy 3054 AAGAAACAATGAATGCTTTTTTCAAACTTTATGACAAAGTATTTTCAAGTCTGTGTTTC 3113  
Db 3130 AAGAAACAATGAATGCTTTTTTCAAACTTTATGACAAAGTATTTTCAAGTCTGTGTTTC 3189

Qy 3114 AAAAAATATATTTCAATGTACCTGTGATCCAGCAAGAGGAGTTCAGTCAAGAGTCACTAC 3173  
Db 3190 AAAAAATATATTTCAATGTACCTGTGATCCAGCAAGAGGAGTTCAGTCAAGAGTCACTAC 3249  
Qy 3174 AACTGATTAGTGTGTTTAGAGAAATGAGAAATGGAACAGTGAAGAAATGGAGCCATATTTCC 3233  
Db 3250 AACTGATTAGTGTGTTTAGAGAAATGAGAAATGGAACAGTGAAGAAATGGAGCCATATTTCC 3309  
Qy 3234 ATGACTTCCCTTCTTAAACAGAAAGCAAGAGGCAAGAGGCTGGCTCTACATCACTC 3293  
Db 3310 ATGACTTCCCTTCTTAAACAGAAAGCAAGAGGCAAGAGGCTGGCTCTACATCACTC 3369  
Qy 3294 TCACCTTCCAAATCTTTGTGGAAGTGCATCTACTTGCAGAACCAAAATTAATCTTACTTTCCA 3353  
Db 3370 TCACCTTCCAAATCTTTGTGGAAGTGCATCTACTTGCAGAACCAAAATTAATCTTACTTTCCA 3429  
Qy 3354 AGTCTTGGCTGCTTGCAGGTGGAACCTCCAGCTGCAAGGAGGTTAGGGAATAGAGTCTT 3413  
Db 3430 AGTCTTGGCTGCTTGCAGGTGGAACCTCCAGCTGCAAGGAGGTTAGGGAATAGAGTCTT 3489  
Qy 3414 TTTTTTAAAGCTTCTCAGACCTTCTAGGGAACAGAAATTTGGGTGAGCAATCTGCAATTT 3473  
Db 3490 TTTTTTAAAGCTTCTCAGACCTTCTAGGGAACAGAAATTTGGGTGAGCAATCTGCAATTT 3549  
Qy 3474 CTACTACAGGCATTTGAGACCCAGTTAGATTTATGAAATATTTATAGAGAGTTTATGAACACTT 3533  
Db 3550 CTACTACAGGCATTTGAGACCCAGTTAGATTTATGAAATATTTATAGAGAGTTTATGAACACTT 3609  
Qy 3534 AAATATGATAGTGGTATGACATTTGATAGAACATGGGATACCTTTAGAGTGAATTTGAC 3593  
Db 3610 AAATATGATAGTGGTATGACATTTGATAGAACATGGGATACCTTTAGAGTGAATTTGAC 3669  
Qy 3594 AGGGCATATTTAGTTGATGAAATGGAGTCAATTTGAGTCTTTTATAGCCATGATCATAAT 3653  
Db 3670 AGGGCATATTTAGTTGATGAAATGGAGTCAATTTGAGTCTTTTATAGCCATGATCATAAT 3729  
Qy 3654 TACCAAGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGGAATATAATGGAC 3713  
Db 3730 TACCAAGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGGAATATAATGGAC 3789  
Qy 3714 AGATTAATATTTGTTCTGTGATGCCCAAACTCCCTTTCTAAGGAAGATGCGCCTACTAT 3773  
Db 3790 AGATTAATATTTGTTCTGTGATGCCCAAACTCCCTTTCTAAGGAAGATGCGCCTACTAT 3849  
Qy 3774 AGCAGTTTTTATATTTGTCATTTTATGAATATAATGAATGAAGAGTCTTGGTACCTCCTG 3833  
Db 3850 AGCAGTTTTTATATTTGTCATTTTATGAATATAATGAATGAAGAGTCTTGGTACCTCCTG 3908  
Qy 3834 TCTTTACAAATATTTGGGTGTTGTCAGATATTTTCCCTTTTAAACCMWTTCCCAATTCGGG 3893  
Db 3909 TCTTTACAAATATTTG---TGTGTGTCAGTATTTTCCCTTTTAAACCAATTCCAATCGGTG 3964  
Qy 3894 TGTGATAGTGGATGTTTCCATTTGGGTTTTTAAATTTGTATATATCCCTGATAGTATAATGG 3953  
Db 3965 TGTAGTATG---TTTCATTTTGGTTTTTAAATTTGTATATATCCCTGATAGTATAATGG 4019  
Qy 3954 GTCATAGAAATTTCTTTATACATTTAGATGCAAGTCTCTTGYCGGATATACGATTTGAGA 4013  
Db 4020 GTCATAGAAATTTCTTTATACATTTAGATGCAAGTCTCTTGTGCGGATATATGATTTGAGA 4079  
Qy 4014 TATTACACCTAGTCTGTGCTTGACTGTTTCTTTTATGTCTTTTGTATGATAGAAAGTTTT 4073  
Db 4080 TATTACACCTAGTCTGTGCTTGACTGTTTCTTTTATGTCTTTTGTATGATAGAAAGTTTT 4139  
Qy 4074 AAATTTTGAAGGTCAAAATTTATTTTCTTTTGTGATATTTTCTCTCCAAATTT 4133  
Db 4140 AAATTTTGAAGGTCAAAATTTATTTTCTTTTGTGATATTTTCTCTCCAAATTT 4199  
Qy 4134 AACCCCAAGATTTGAGATTTCTGCTCTATATATAAACTTTATATTTTTTATATTTTGTGA 4193  
Db 4200 AACCCCAAGATTTGAGATTTCTGCTCTATATATAAACTTTATATTTTTTATATTTTGTGA 4259



QY 4194 TCTACCTTGAATTGATATGATGTTGTGAATTATGATCAGGGTCTCTTTTTCGCCCAT 4253  
DB |||||  
DB 4260 TCTACCTTGAATTGATATGATGTTGTGAATTATGATCAGGGTCTCTTTTTCGCCCAT 4319  
QY 4254 ACAAGTATCCAGTCATTTGTAACACTGTTTATTGAAAGAAATTATCTTCTCATTAATAATT 4313  
DB |||||  
DB 4320 ACAAGTATCCAGTCATTTGTAACACTGTTTATTGAAAGAAATTATCTTCTCATTAATAATT 4379  
QY 4314 ACCTTCCCAATTAGTAAAAAATCAATTAACCAT 4346  
DB |||||  
DB 4380 ACCTTCCCAATTAGTAAAAAATCAATTAACCAT 4412

RESULT 2  
AAZ89047  
ID AAZ89047 standard; DNA; 4386 BP.  
AC AC  
XX AAZ89047;  
XX 01-JUN-2000 (first entry)  
XX Human nibrin DNA.  
DE  
XX Nibrin; human; DNA double strand break repair protein; diagnosis;  
KW therapy; Nijmegen Breakage Syndrome; gene therapy; ds.  
XX  
XX Homo sapiens.  
XX  
XX  
FH Key Location/Qualifiers  
FT CDS 26..2290  
FT /\*tag= a  
FT /product= "Nibrin"XX  
PN DE19818680-C1.  
XX  
XX 09-MAR-2000.  
XX 27-APR-1998; 98DE-01018680.  
XX  
XX 27-APR-1998; 98DE-01018680.  
PR (UYBE ) UNIV BERLIN HUMBOLDT.  
XX  
XX WPI: 2000-196117/18.  
DR P-PSDB; AAY51669.  
XX  
XX  
PT A DNA double strand break repair protein, Nibrin, and related DNA useful  
PT for diagnosis and therapy of Nijmegen Breakage Syndrome and other  
PT diseases influenced by DNA-double-strand break repair.  
XX  
XX  
PS Claim 2a; Fig 2; 32pp; German.  
XX  
XX This invention describes a novel DNA double strand break repair protein,  
CC Nibrin. Nibrin and DNA encoding it are useful for diagnosis and/or  
CC therapy of diseases influenced by repair of DNA-double strand breaks, in  
CC particular Nijmegen Breakage Syndrome. The product of the invention has  
CC applications in gene therapy. This sequence encodes the nibrin protein  
CC described in the invention  
XX  
SQ Sequence 4386 BP; 1467 A; 744 C; 845 G; 1330 T; 0 U; 0 Other;

Query Match 93.2%; Score 4105.2; DB 3; Length 4386;  
Best Local Similarity 97.1%; Pred. No. 0;  
Matches 4270; Conservative 7; Mismatches 30; Indels 91; Gaps 6;  
QY 28 GCCCAGCCCTGAGGAGCCGACCGATGTGGAACCTGCTGCCGCCGCCGCCGCCGCCGAGG 87  
DB 1 GCCCAGCCCTGAGGAGCCGACCGATGTGGAACCTGCTGCCGCCGCCGCCGCCGCCGAGG 60  
QY 88 AGGAGAACCATACAGACTTTTGTGCTGGCTGTGAGTACGTTGTTGGAAGAAAACTGTGC 147  
DB 61 AGGAGAACCATACAGACTTTTGTGCTGGCTGTGAGTACGTTGTTGGAAGAAAACTGTGC 120

QY 148 CATTCTAATTTGAAATGATCAGTCGATCAGCGGAAATCATGCTGTGTTAACTGCTAACTT 207  
DB |||||  
DB 121 CATTCTGATTTGAAATGATCAGTCGATCAGCGGAAATCATGCTGTGTTAACTGCTAACTT 180  
QY 208 TTCTGTAAACCAACCTGAGTCACAAAGATGAAATCCCTGTATTTGACATTTAAAGATAATTC 267  
DB |||||  
DB 181 TTCTGTAAACCAACCTGAGTCACAAAGATGAAATCCCTGTATTTGACATTTAAAGATAATTC 240  
QY 268 TAAGTATGGTACCTTTGTTAATGAGGAAAAATGAGAAATGGCTTTTCCCGAACTTTCGA 327  
DB |||||  
DB 241 TAAGTATGGTACCTTTGTTAATGAGGAAAAATGAGAAATGGCTTTTCCCGAACTTTCGA 300  
QY 328 GTCGGGGGATGGTATTACTTTTGGAGTGTCTTGGAGTAAATTCAGAAATAGAGTATGAGCC 387  
DB |||||  
DB 301 GTCGGGGGATGGTATTACTTTTGGAGTGTCTTGGAGTAAATTCAGAAATAGAGTATGAGCC 360  
QY 388 TTTGGTTGCATGCTCTTCTTGTATTAGATGCTCTGCGGAAAACTGCTTTAAATCAAGCTAT 447  
DB |||||  
DB 361 TTTGGTTGCATGCTCTTCTTGTATTAGATGCTCTGCGGAAAACTGCTTTAAATCAAGCTAT 420  
QY 448 ATTGCAACTTTGGAGGATTTACTGTAAACAATTTGGACAGAAGATGCACCTTGTGTCAT 507  
DB |||||  
DB 421 ATTGCAACTTTGGAGGATTTACTGTAAACAATTTGGACAGAAGATGCACCTTGTGTCAT 480  
QY 508 GGTATCAGTGAAGTTTACCATTTAAACAATATGTCACCTCATTTTGTGACGTCCTCAATTGT 567  
DB |||||  
DB 481 GGTATCAGTGAAGTTTACCATTTAAACAATATGTCACCTCATTTTGTGACGTCCTCAATTGT 540  
QY 568 AAGCCAGATAATTTTACTGAATTCCTGAAAGCAGTTTCAGTCCAAGAGAGCCTCCACA 627  
DB |||||  
DB 541 AAGCCAGATAATTTTACTGAATTCCTGAAAGCAGTTTCAGTCCAAGAGAGCCTCCACA 600  
QY 628 AATTGAAAGTTTTCACCCACCTCTTGATGAACCATCTATTGGAAGTAAATTTGTGATCT 687  
DB |||||  
DB 601 AATTGAAAGCTTTTACCCACCTCTTGATGAACCATCTATTGGAAGTAAATTTGTGATCT 660  
QY 688 GTCAGGACGCGAGGAAAGAAACAAATCTTCAAAGGGAAAAACATTTATTTTGAATGC 747  
DB |||||  
DB 661 GTCAGGACGCGAGGAAAGAAACAAATCTTCAAAGGGAAAAACATTTATTTTGAATGC 720  
QY 748 CAAACAGCATTAAGAAATTTAGTTCGCGAGTTGCTTTTGGAGTGGGGAGCTAGGTTGAT 807  
DB |||||  
DB 721 CAAACAGCATTAAGAAATTTAGTTCGCGAGTTGCTTTTGGAGTGGGGAGCTAGGTTGAT 780  
QY 808 AACAGAGAGATGAAGAGAACATAATTTCTTTTGGCTCGGGAAACGTGTGTTGTTGA 867  
DB |||||  
DB 781 AACAGAGAGATGAAGAGAACATAATTTCTTTTGGCTCGGGAAACGTGTGTTGTTGA 840  
QY 868 TACAGGAATAACAACTCACAGACCTTAAATTCCTGACTGTGAGAAAGAAATGGATTTCAGTC 927  
DB |||||  
DB 841 TACAGGAATAACAACTCACAGACCTTAAATTCCTGACTGTGAGAAAGAAATGGATTTCAGTC 900  
QY 928 AATAATGATATGCTCCAAAGCAAGGTCCTTAGACCTATTCTGTAAGCAGAAAAATTTGGATT 987  
DB |||||  
DB 901 AATAATGATATGCTCCAAAGCAAGGTCCTTAGACCTATTCTGTAAGCAGAAAAATTTGGATT 960  
QY 988 GCGGTGATTTTCATGACTACAAAGAAATTTACTGTGATCCTCAGGSCCATCCAGTACAGG 1047  
DB |||||  
DB 961 GCGGTGATTTTCATGACTACAAAGAAATTTACTGTGATCCTCAGGSCCATCCAGTACAGG 1020  
QY 1048 ATTAAGACAACTCCAGGACCAAGCCTTTTCAAGCGGTGTGAGTTGATGAAAAACT 1107  
DB |||||  
DB 1021 ATTAAGACAACTCCAGGACCAAGCCTTTTCAAGCGGTGTGAGTTGATGAAAAACT 1080  
QY 1108 AATGCAAGCGCCGAGTGAACTATCAACATATAGTAGTGAACAGAAATCGAGCAAGC 1167  
DB |||||  
DB 1081 AATGCAAGCGCCGAGTGAACTATCAACATATAGTAGTGAACAGAAATCGAGCAAGC 1140  
QY 1168 AGATACATGGGATTTGAGTGAAGGCCAAAGAAATCAAGTCTCCAAAAATGGAACAAA 1227  
DB |||||  
DB 1141 AGATACATGGGATTTGAGTGAAGGCCAAAGAAATCAAGTCTCCAAAAATGGAACAAA 1200  
QY 1228 ATTCAGAAATGCTTTTCAAGAGCGCACCCACTGTAAAGGAGTCTCTGCAAAAACAAGCTCTAA 1287



Db 1201 ATTGAGAAATGCTTTTCAAGATGACCCACTGTAAGGAGTCTCTGCAAAACAAAGCTCTAA 1260  
Qy 1288 TAATAATAGTATGGTATCAAAATACTTTGGCTAAGATGAGAAATCCCAAATCTACAGCTTTC 1347  
Db 1261 TAATAATAGTATGGTATCAAAATACTTTGGCTAAGATGAGAAATCCCAAATCTACAGCTTTC 1320  
Qy 1348 ACCAACTAAATGCCAAAGTATAATAAAGTAAAGATAGGCTTCTCAGCAGCAGCAGAC 1407  
Db 1321 ACCAACTAAATGCCAAAGTATAATAAAGTAAAGATAGGCTTCTCAGCAGCAGCAGAC 1380  
Qy 1408 CAACTCCATCAGAACTACTTTTCAGCGCTCTACCAAAAAGGAAAGGATGAGAAAA 1467  
Db 1381 CAACTCCATCAGAACTACTTTTCAGCGCTCTACCAAAAAGGAAAGGATGAGAAAA 1440  
Qy 1468 TCAAGAAATGCTTTTCATGCAAAATCAGCAAGATAGAAAAGCTGTGTTCTCTTTTAGAACA 1527  
Db 1441 TCAAGAAATGCTTTTCATGCAAAATCAGCAAGATAGAAAAGCTGTGTTCTCTTTTAGAACA 1500  
Qy 1528 AACAACTGCTACACCTCTATGTTGGAAAAATAAGGAGCAGCATCTATCTCAGAAATGA 1587  
Db 1501 AACAACTGCTACACCTCTATGTTGGAAAAATAAGGAGCAGCATCTATCTCAGAAATGA 1560  
Qy 1588 GCCTGTGGACAACTCAGACAAATTAATTTTACAGATACAGATTTAAATCTATTGT 1647  
Db 1561 GCCTGTGGACAACTCAGACAAATTAATTTTACAGATACAGATTTAAATCTATTGT 1620  
Qy 1648 GAAAAATTCGCGAGTAAATCTCATGCTGCAGAAAAAGCTAAGATCAAAATAAAAAAGGGA 1707  
Db 1621 GAAAAATTCGCGAGTAAATCTCATGCTGCAGAAAAAGCTAAGATCAAAATAAAAAAGGGA 1680  
Qy 1708 AATGGATGATGTGGCCATAGAAGATGAAGTATTGGAACAGTATTCAAGGACACAAAACC 1767  
Db 1681 AATGGATGATGTGGCCATAGAAGATGAAGTATTGGAACAGTATTCAAGGACACAAAACC 1740  
Qy 1768 AGATTAGAAATTTGATGTGAAGTTCAAAAACAGGAGGAAGATGTCNAATTTAGAAAAAG 1827  
Db 1741 AGATTAGAAATTTGATGTGAAGTTCAAAAACAGGAGGAAGATGTCNAATTTAGAAAAAG 1800  
Qy 1828 GCCAAGATGGATATAGAAACAAATGACACTTTTCAGTGATGAAGCAGTACCAGAAAGTAG 1887  
Db 1801 GCCAAGATGGATATAGAAACAAATGACACTTTTCAGTGATGAAGCAGTACCAGAAAGTAG 1860  
Qy 1888 CAAAATATCTCAAGAAAAATGAAATTTGGGAAGAAACGTGAACTCAAGGAAGACTCACTATG 1947  
Db 1861 CAAAATATCTCAAGAAAAATGAAATTTGGGAAGAAACGTGAACTCAAGGAAGACTCACTATG 1920  
Qy 1948 GTCAGCTAAAGAAATATCTAACAAATGACAAACTTCAGGATGATAGTGAGATGCTTCCAAA 2007  
Db 1921 GTCAGCTAAAGAAATATCTAACCAATGACAACTTCAGGATGATAGTGAGATGCTTCCAAA 1980  
Qy 2008 AAAGCTGTTATTGCACTGAAATTTAGATCAGCTGGTGATTTAAAACTCTACTTCCAGAAATCC 2067  
Db 1981 AAAGCTGTTATTGCACTGAAATTTAGATCAGTGGTGATTTAAAACTCTACTTCCAGAAATCC 2040  
Qy 2068 GTCTGCATAAATGATGATTATGGTCAACTAAAAATTTTCAAGAAATTTCAAAAAGGTCA 2127  
Db 2041 ATCTGGCATAAATGATGATTATGGTCAACTAAAAATTTTCAAGAAATTTCAAAAAGGTCA 2100  
Qy 2128 ATATCTCGAGCAGGAAAACTTCCACACATCATTTGGAGGATCAGATCTAATAGCTCATCA 2187  
Db 2101 ATATCTCGAGCAGGAAAACTTCCACACATCATTTGGAGGATCAGATCTAATAGCTCATCA 2160  
Qy 2188 TGCTCGAAAGAAATACAGAACTACAGAGTGGCTTAAGCAGGAAATGAGGTACAAAATCA 2247  
Db 2161 TGCTCGAAAGAAATACAGAACTACAGAGTGGCTTAAGCAGGAAATGAGGTACAAAATCA 2220  
Qy 2248 ACATGCAAAAGAAAGAGTCTCTGCTGATGATCTTTTATAGATACAACTCTTATTTAAAAAG 2307  
Db 2221 ACATGCAAAAGAAAGAGTCTCTGCTGATGATCTTTTATAGATACAACTCTTATTTAAAAAG 2280  
Qy 2308 GAGAAATTAATCTGAGGATTTTAAAAAGAGCCATGGAAAACTTCTCTAGTAAAGCTCTAC 2367

Db 2281 GAGAAGATTAATGAGGATTTTAAAAAGAGCCATGGAAAAAATCTTCTAGTAAAGCATCTAC 2340  
Qy 2368 TTCAAGCCCAACAGGTTATATGAATATATAGTGTATAGAGCGATTTAAAGTTACAATGTT 2427  
Db 2341 TTCAAGCCCAACAGGTTATATGAATATATAGTGTATAGAGCGATTTAAAGTTACAATGTT 2400  
Qy 2428 TTATGCGCTAAATTTATTAATAAATATGACAAAACTTTGATTTCTTTTGTATGTAAACAAT 2487  
Db 2401 TTATGCGCTAAATTTATTAATAAATATGACAAAACTTTGATTTCTTTTGTATGTAAACAAT 2460  
Qy 2488 TGTGTTGCTGTTTTCAGGCTTTGTCATGTCATCTTTTTCATTTTAAATGTTGTTTTC 2547  
Db 2461 TGTGTTGCTGTTTTCAGGCTTTGTCATGTCATCTTTTTCATTTTAAATGTTGTTTTC 2520  
Qy 2548 TTATTAATAATAGTAAATATATAGTACACAGTTCAAAATTTCTAAATRTACGTAAAGTAAAGGAC 2607  
Db 2521 TTATTAATAATAGTAAATATATAGTACACAGTTCAAAATTTCTAAATRTACGTAAAGTAAAGGAC 2579  
Qy 2608 TAAAGTCACCTTCACCAATTTGCTCCTAGTACT-----TATTTTAA 2640  
Db 2580 TAAAGTCACCTTCACCAATTTGCTCCTAGTACTTTGCTTCCCTCAGAAAAAATTCATGA 2639  
Qy 2641 -----TATTTTAA 2648  
Db 2640 TACTCATTTCTTAAGAAATCTTTCCAGGGATTTTGTAGTCTTATTCAAATTTCTTATTTTAA 2699  
Qy 2649 AATAAATTTCTCACAATAATGATAGATAACATATGAGTGTGTTCTACACCTTGTCTTTTAA 2708  
Db 2700 AATAAATTTCTCACAATAATGATAGATAACATATGAGTGTGTTCTACACCTTGTCTTTTAA 2759  
Qy 2709 CTTAGTAAGATTAATAAATTTATAGGAATATCAATATAATGTTTTTAAATATTTTTTCTTTTC 2768  
Db 2760 CTTAGT-AGATTAATAAATTTATAGGAATATCAATATAATGTTTTTAAATATTTTTTCTTTTC 2818  
Qy 2769 CATTATGCTGTAGTCTTACCTAAACCTCTGATCAABACAAATGCGCTTCAGTGTGCGCA 2828  
Db 2819 CATTATGCTGTAGTCTTACCTAAACCTCTGATCAABACAAATGCGCTTCAGTGTGCGCA 2878  
Qy 2829 GATGTCACTACATGTTTATTCTAGTACTAGAAACCTGAAGACCATGTGGAGACTTCATCAA 2888  
Db 2879 GATGTCACTACATGTTTATTCTAGTACTAGAAACCTGAAGACCATGTGGAGACTTCATCAA 2938  
Qy 2889 ACATGGGTTTGTGTTTACCAGAAATGGAAGACCTGTACCCCTTTTGTGGTGTCTTACTG 2948  
Db 2939 ACATGGGTTTGTGTTTACCAGAAATGGAAGACCTGTACCCCTTTTGTGGTGTCTTACTG 2998  
Qy 2949 AGCTGGGTGGGTGCTGTTTTCAGGCTTATTTAGAGTCTTAGTCTTCTCTATTAAGTA 3008  
Db 2999 AGCTGGGTGGGTGCTGTTTTCAGGCTTATTTAGAGTCTTAGTCTTCTCTATTAAGTA 3058  
Qy 3009 GAAATGGTGAGATTTGTTTCTTTTCTACCKTAAAGGGAGATGGTAAAGAAACAAATGAAATG 3068  
Db 3059 GAAATGGTGAGATTTGTTTCTTTTCTACCTTAAAGGGAGATGGTAAAGAAACAAATGAAATG 3118  
Qy 3069 TCTTTTTTCAAACTTTTATTTGA CAAGTATTTTCAAGTCTGTGTTTCAAAAATATATTTATG 3128  
Db 3119 TCTTTTTTCAAACTTTTATTTGA CAAGTATTTTCAAGTCTGTGTTTCAAAAATATATTTATG 3178  
Qy 3129 TACCTGTGATCCAGCAAGAGGAGTTCAGTCAAGAGTCACTACAACTGATTAGTGTGTT 3188  
Db 3179 TACCTGTGATCCAGCAAGAGGAGTTCAGTCAAGAGTCACTACAACTGATTAGTGTGTT 3238  
Qy 3189 TAGAAGATGAGAAATGGAAACAGTGAAGGATGGAGGCAATTTTCCATGACTTCCCTTTGTA 3248  
Db 3239 TAGAAGATGAGAAATGGAAACAGTGAAGGATGGAGGCAATTTTCCATGACTTCCCTTTGTA 3298  
Qy 3249 AACAGAGCAA CAGAGGGA CAAGGCTGGGCTCTTACATCACTCTCACCTTCCAAATCT 3308  
Db 3299 AACAGAGCAA CAGAGGGA CAAGGCTGGGCTCTTACATCACTCTCACCTTCCAAATCT 3358  
Qy 3309 TGTGGAGTGCATCTACTTTCAGCAACCAAAATTAATTTACTTCTCCAGTTCTGGCTGCTG 3368  
Db 3359 TGTGGAGTGCATCTACTTTCAGCAACCAAAATTAATTTACTTCTCCAGTTCTGGCTGCTG 3418



```
QY 3369 CAGGTGGAACCTCAGCTGCAAGGGAGCTTAGGGAATGAAGGCTCTTTTAAAGCTTCT 3428
DB 3419 CAGGTGGAACCTCAGCTGCAAGGGAGCTTAGGGAATGAAGGCTCTTTTAAAGCTTCT 3478
QY 3429 CAGCCCTTCCTAGGGAACAGAAATGGGTGAGCCAAATCTGCAATTTCTACTACAGGCATTTG 3488
DB 3479 CAGCCCTTCCTAGGGAACAGAAATGGGTGAGCCAAATCTGCAATTTCTACTACAGGCATTTG 3538
QY 3489 AGACCAAGTAGATTATGAAATATATATAGAGAGTTATGAACACTTAAATATATGATAGTGG 3548
DB 3539 AGACCAAGTAGATTATGAAATATATATAGAGAGTTATGAACACTTAAATATATGATAGTGG 3598
QY 3549 TATGACATTCGATAGAACATGGGATACCTTTAGAGTAGAATTCACAGGGCATATAGTTG 3608
DB 3599 TATGACATTCGATAGAACATGGGATACCTTTAGAGTAGAATTCACAGGGCATATAGTTG 3658
QY 3609 ATGAAATGGAGTCATTTGAGTCTTAAATAGCCATGTATCAATTAACCAAGTGAAGCTG 3668
DB 3659 ATGAAATGGAGTCATTTGAGTCTTAAATAGCCATGTATCAATTAACCAAGTGAAGCTG 3718
QY 3669 GTGGAACATATGCTCTCCATTTTACAGTTAAGGAATATATGAGACAGATTAATATTTGTTT 3728
DB 3719 GTGGAACATATGCTCTCCATTTTACAGTTAAGGAATATATGAGACAGATTAATATTTGTTT 3778
QY 3729 TCTGTCTATGCCACATCCCTTCTTAAGGAAGTCTGCTGTAATCTCTGCTTTTACAAATATTTG 3788
DB 3779 TCTGTCTATGCCACATCCCTTCTTAAGGAAGTCTGCTGTAATCTCTGCTTTTACAAATATTTG 3838
QY 3789 TGTCAATTTATGATATATATGATAGAGAGTCTGCTGTAATCTCTGCTTTTACAAATATTTG 3848
DB 3839 TGTCAATTTATGATATATATGATAGAGAGTCTGCTGTAATCTCTGCTTTTACAAATATTTG 3897
QY 3849 GGTGTTGTCAGATATTTTCCCTTTTAAACCMWTCCTCAATTCGGGTGTAGTGAATGT 3908
DB 3898 G-----TGTTGTGAGTATTTTCTTTTAAACCATTCGAATCGGTGTAGTGAATG----- 3948
QY 3909 TTCCATTTGGTGTATTTGATATATCCCTGATAGCTATATTTGGTGTATAGAAATTCCT 3968
DB 3949 TTTCAATTTGGTGTATTTGATATATCCCTGATAGCTATATTTGGTGTATAGAAATTCCT 4008
QY 3969 TATACATTTAGATCAAGTCTCTGCGGATATACGTATTTGATATATTTACACCTAGTCT 4028
DB 4009 TATACATTTAGATCAAGTCTCTCTCGGATATATGATATGATATATACACCTAGTCT 4068
QY 4029 GTGGCTGTAGTGTCTTTATATGCTTTTATGATAGAAATTTTAAATTTTGAACAAGT 4088
DB 4069 GTGGCTGTAGTGTCTTTATGCTTTTATGATAGAAATTTTAAATTTTGAACAAGT 4128
QY 4089 CAAATTTATTTTCTTTTGTGATATTTTCTCTCCAAATTTAAACCCCAAGATTTCA 4148
DB 4129 CAAATTTATTTTCTTTTGTGATATTTTCTCTCCAAATTTAAACCCCAAGATTTCA 4188
QY 4149 GATATTTCTGCTATATATATAAACTTTATATTTTATATTTTGTGATCTACCTTTGAATTGA 4208
DB 4189 GATATTTCTGCTATATATAAACTTTATATTTTATATTTTGTGATCTACCTTTGAATTGA 4248
QY 4209 TATGATTTGTGAATATATGATATCAGGGTCTTTTTCCTCCCAATCAAGATATCCAGTCA 4268
DB 4249 TATGATTTGTGAATATATGATATCAGGGTCTTTTTCCTCCCAATCAAGATATCCAGTCA 4308
QY 4269 TTGTACACTGTTTATTTGAAGAAATATCTTTCTTCATTAATTTACCTTGGCAATTTAGT 4328
DB 4309 TTGTACACTGTTTATTTGAAGAAATATCTTTCTTCATTAATTTACCTTGGCAATTTAGT 4368
QY 4329 AAAAAATCAATTAACCAT 4346
DB 4369 AAAAAATCAATTAACCAT 4386
```

RESULT 3  
AAZ34997  
ID AAZ34997 standard; cdna; 4386 BP.

```
XX AAZ34997;
XX AC
XX DT 28-FEB-2000 (first entry)
XX DE NBS1 gene associated with Nijmegen breakage syndrome.
XX KW NBS1 gene; nibrin; Nijmegen breakage syndrome; diagnosis; human;
XX KW gene therapy; cancer; microcephaly; mental retardation;
XX KW primary ovarian failure; ss.
XX OS Homo sapiens.
XX FH Location/Qualifiers
XX CDS 26..2290
XX FT /*tag= a
XX FT polyA_signal 2420..2425
XX FT /*tag= b
XX FT repeat_region 3853..4386
XX FT /*tag= c
XX FT /note= "LMC/D genome wide repeat"
XX PN WO9955716-A1.
XX XX 04-NOV-1999.
XX XX 27-APR-1999; 99WO-US009036.
XX XX 27-APR-1999; 98US-0083269P.
XX XX (VIRG-) VIRGINIA MASON RES CENT.
XX XX Concannon PJ, Vissinga CS, Cerosaletti KM, Varon R, Sperling K;
XX XX Reis A;
XX XX WPI; 2000-062015/05.
XX XX P-PSDB; AAY32373.
XX XX Novel gene useful for detecting mutations or polymorphisms, and
XX XX diagnosing certain pathological conditions in Nijmegen Breakage syndrome
XX XX patients.
XX XX Claim 2; Fig 2A-C; 58pp; English.
XX XX This is the nucleotide sequence of cDNA for the NBS1 gene that is
XX XX associated with the Nijmegen breakage syndrome (NBS). It includes a
XX XX coding region for a 754-amino acid protein, nibrin (see AAY32373). The
XX XX gene maps to a 1 cm region on chromosome 8q21. The invention is based on
XX XX the discovery that the gene contains mutations in all NBS patients. These
XX XX mutations include deletions and insertions that result in frameshift, as
XX XX well as point mutations. Specific mutations associated with the NBS
XX XX phenotype include 657del5, 698del4, 835del4, 842insT, 1142delC, 976C>T,
XX XX 681delT and 900del25. Polymorphisms include 553 G/C, 1197 T/C, 2016 A/G,
XX XX 102 G/A, IVS 5+9 T/C, IVS5+51delT, IVS9+18 C/T and IVS-7A/G. It is an
XX XX object of the invention to detect a mutation or polymorphism in NBS
XX XX patients, and hence to diagnose a predisposition to a pathological
XX XX condition such as cancer, microcephaly, mental retardation, and primary
XX XX ovarian failure, based on detection of a mutation in the NBS1 gene. It is
XX XX also an object of the invention to treat NBS by replacing the mutated
XX XX gene in a NBS patient by gene therapy. Recombinant vectors, genetically
XX XX engineered host cells, a method for producing nibrin polypeptide, an
XX XX antibody that specifically binds to the polypeptide, and a method for
XX XX diagnosing NBS are claimed. Primers (see AAZ34998-5035) used for
XX XX detection a mutation in NBS1 are also claimed
XX SQ Sequence 4386 BP; 1467 A; 747 C; 841 G; 1331 T; 0 U; 0 Other;
Query Match 93.2%; Score 4103.6; DB 3; Length 4386;
Best Local Similarity 97.1%; Pred No. 0;
Matches 4269; Conservative 7; Mismatches 31; Indels 91; Gaps 6;
QY 28 GCCCCAGCCCTTAGAGAGCCGACCATGTGGAATCTGTGCCGCCGCCGCCGCCGAGG 87
|||||
```



Db 1 GCCCCAGCCCTGAGGAGCCGGA CCGATGTGGAACCTGCTGCCCGCCGCGCCGCGCAGG 60  
Qy 88 AGGAGAACCATACAGACTTTTGACTGGCGTTGAGTAGCTGTTGGTGGAGGAAAAAAGCTGTC 147  
Db 61 AGGAGAACCATACAGACTTTTGACTGGCGTTGAGTAGCTGTTGGTGGAGGAAAAAAGCTGTC 120  
Qy 148 CATTCTAAATTGAAAATGATCAGTCGATCAGCCGAAATCATGCTGTGTAACTGCTAACTT 207  
Db 121 CATTCTGATTGAAAATGATCAGTCGATCAGCCGAAATCATGCTGTGTAACTGCTAACTT 180  
Qy 208 TTCTGTAAACCAACCTCAGTCAAAACAGATGAAATCCCTGTATTGACATTAAGAAATATTC 267  
Db 181 TTCTGTAAACCAACCTCAGTCAAAACAGATGAAATCCCTGTATTGACATTAAGAAATATTC 240  
Qy 268 TAAGTAGGTGACCTTTGTTAATGAGGAAAAATGCAAGATGGCTTTTCCCGAACTTTGAA 327  
Db 241 TAAGTAGGTGACCTTTGTTAATGAGGAAAAATGCAAGATGGCTTTTCCCGAACTTTGAA 300  
Qy 328 GTCGGGGATGGTATTACTTTTGGAGTGTGTTGGAAGTAAATTCAGAAATGAGTATGAGCC 387  
Db 301 GTCGGGGATGGTATTACTTTTGGAGTGTGTTGGAAGTAAATTCAGAAATGAGTATGAGCC 360  
Qy 388 TTTGGTGTGATGCTCTCTGTTTGTAGATGCTCTGGGAAAACTGCTTTAAATCAAGCTAT 447  
Db 361 TTTGGTGTGATGCTCTCTGTTTGTAGATGCTCTGGGAAAACTGCTTTAAATCAAGCTAT 420  
Qy 448 ATTGCAACTTTGGAGGATTTACTGTAAACAAATTGGACAGAGAAATGCACCTTGTCAT 507  
Db 421 ATTGCAACTTTGGAGGATTTACTGTAAACAAATTGGACAGAGAAATGCACCTTGTCAT 480  
Qy 508 GGTATCAGTGAAGTTACCAATTAACAAATATGTGCACCTCATTTTGGACGTCCTCAATTTGT 567  
Db 481 GGTATCAGTGAAGTTACCAATTAACAAATATGTGCACCTCATTTTGGACGTCCTCAATTTGT 540  
Qy 568 AAGCCAGATATTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAGAACGAGCTCCACA 627  
Db 541 AAGCCAGATATTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAGAACGAGCTCCACA 600  
Qy 628 AATTGAAAGTTTATCCACCTCTGATGAACCATCTATTGGAAGTAAAAATGTTGATCT 687  
Db 601 AATTGAAAGTTTATCCACCTCTGATGAACCATCTATTGGAAGTAAAAATGTTGATCT 660  
Qy 688 GTCAGACGCGAGGAAGAAAAACAATCTTCAAAGGAAAAACAATTTATATTTTGAATGC 747  
Db 661 GTCAGACGCGAGGAAGAAAAACAATCTTCAAAGGAAAAACAATTTATATTTTGAATGC 720  
Qy 748 CAACACGATTAAGAAATGAGTTCCGAGTTGCTTTGGAGGTGGGAGAGCTAGGTTGAT 807  
Db 721 CAACACGATTAAGAAATGAGTTCCGAGTTGCTTTGGAGGTGGGAGAGCTAGGTTGAT 780  
Qy 808 AACAGAGAGATGAAGAGAACATAATTTCTTTTGGCTCCGGGAACGTTGTTGTA 867  
Db 781 AACAGAGAGATGAAGAGAACATAATTTCTTTTGGCTCCGGGAACGTTGTTGTA 840  
Qy 868 TACAGGAATAACAAACTCAAGACCTTAATTCCTGATCTGTCAAGAGAAATGGATTGAGTC 927  
Db 841 TACAGGAATAACAAACTCAAGACCTTAATTCCTGATCTGTCAAGAGAAATGGATTGAGTC 900  
Qy 928 AATAATGGATATGCTCCAAAGGCAAGCTTTAGACCTATTCCTGAAGCAGAAATGGATT 987  
Db 901 AATAATGGATATGCTCCAAAGGCAAGCTTTAGACCTATTCCTGAAGCAGAAATGGATT 960  
Qy 988 GGGGGTGAATTTTATGACTACAAAGAAATTTACTGTGATCTCAGGGCCATCCAGTACAGG 1047  
Db 961 GGGGGTGAATTTTATGACTACAAAGAAATTTACTGTGATCTCAGGGCCATCCAGTACAGG 1020  
Qy 1048 ATTAAAGACAACAACTCCAGGACCAAGCCCTTTCAGAGGCGTGTGAGTTGATGAAAAAAGCT 1107  
Db 1021 ATTAAAGACAACAACTCCAGGACCAAGCCCTTTCAGAGGCGTGTGAGTTGATGAAAAAAGCT 1080  
Qy 1108 AATGCCAAGGCCCGGAGTGAACACTACACATACGTAGCTGACACAGAAATCAGAGCAAGC 1167  
Db 1081 AATGCCAAGGCCCGGAGTGAACACTACACATACGTAGCTGACACAGAAATCAGAGCAAGC 1140

Qy 1168 AGATACATGGGATTTTGAAGTGAAGGCGCAAGAAATCAAAGTCTCCAAATGGAACAAA 1227  
Db 1141 AGATACATGGGATTTTGAAGTGAAGGCGCAAGAAATCAAAGTCTCCAAATGGAACAAA 1200  
Qy 1228 ATTCAAGATGCTTTCAAGACGACCCCACTGTAAAGGAGTCTCTGCAAAAACAGCTCTAA 1287  
Db 1201 ATTCAAGATGCTTTCAAGATGACCCCACTGTAAAGGAGTCTCTGCAAAAACAGCTCTAA 1260  
Qy 1288 TAATAATAGTATGATCAATACATTTTGGCTAAAGATGAGAAATCCCAAACTATCAGCTTTC 1347  
Db 1261 TAATAATAGTATGATCAATACATTTTGGCTAAAGATGAGAAATCCCAAACTATCAGCTTTC 1320  
Qy 1348 ACCAACTAAATTTGCCAAGTATAAATAAAGATAGAGTGGCTTTCTCAGCAGCAGCAGAC 1407  
Db 1321 ACCAACTAAATTTGCCAAGTATAAATAAAGATAGAGTGGCTTTCTCAGCAGCAGCAGAC 1380  
Qy 1408 CAACTCCATCAGAAACTACTTTTCAGCCGTCTACCAAAAAAGGAAAGGATGAAGAAA 1467  
Db 1381 CAACTCCATCAGAAACTACTTTTCAGCCGTCTACCAAAAAAGGAAAGGATGAAGAAA 1440  
Qy 1468 TCAAGAAATGCTTTCATGCAAAATCAGCAAGAAATAGAAACGCTCTTGTCTCTTTTGAACA 1527  
Db 1441 TCAAGAAATGCTTTCATGCAAAATCAGCAAGAAATAGAAACGCTCTTGTCTCTTTTGAACA 1500  
Qy 1528 AACACAACTGCTACACCTCATTTGTGAAAAAATAGGAGCAGCATCTATCTGAGAAATGA 1587  
Db 1501 AACACAACTGCTACACCTCATTTGTGAAAAAATAGGAGCAGCATCTATCTGAGAAATGA 1560  
Qy 1588 GCCTGTGGACACAACTCAGACAAATACTTATTATTACAGATACAGATTTAAAATCTATTGT 1647  
Db 1561 GCCTGTGGACACAACTCAGACAAATACTTATTATTACAGATACAGATTTAAAATCTATTGT 1620  
Qy 1648 GAAATTTCTGCCAGTAAATCTCATGCTGCAGAAAGCTAAAGTCAAATAAAAAAGGGA 1707  
Db 1621 GAAATTTCTGCCAGTAAATCTCATGCTGCAGAAAGCTAAAGTCAAATAAAAAAGGGA 1680  
Qy 1708 AATGATGATGTGGCCATAGAGATGAAGTATTGGAACAGATTATTCAAGGACACAAAAACC 1767  
Db 1681 AATGATGATGTGGCCATAGAGATGAAGTATTGGAACAGATTATTCAAGGACACAAAAACC 1740  
Qy 1768 AGAGTTAGAAATGTAGTGAAGTTCAAAAACAGGAGGAAGATGTCAATGTTAGAAAAAG 1827  
Db 1741 AGAGTTAGAAATGTAGTGAAGTTCAAAAACAGGAGGAAGATGTCAATGTTAGAAAAAG 1800  
Qy 1828 GCCAAGGATGGATATAGAAACAAATGACACTTTTCAGTGAATGAAGCAGTACCAAGAAAGTAG 1887  
Db 1801 GCCAAGGATGGATATAGAAACAAATGACACTTTTCAGTGAATGAAGCAGTACCAAGAAAGTAG 1860  
Qy 1888 CAAAATATCTCAAGAAATGAAATTTGGGAAGAAAACGTGAACTCAAGGAAGACTCACTATG 1947  
Db 1861 CAAAATATCTCAAGAAATGAAATTTGGGAAGAAAACGTGAACTCAAGGAAGACTCACTATG 1920  
Qy 1948 GTCAGCTAAAGAAATATCTAAACAATGACAAACTTCAGGAATGATAGTGAATGCTTCCAAA 2007  
Db 1921 GTCAGCTAAAGAAATATCTAAACAATGACAAACTTCAGGAATGATAGTGAATGCTTCCAAA 1980  
Qy 2008 AAGAGCTGTTATGACTGAAATTTAGATCACTGGTGAATTAAGAACTCTACTTCCAGAAATCC 2067  
Db 1981 AAGAGCTGTTATGACTGAAATTTAGATCACTGGTGAATTAAGAACTCTACTTCCAGAAATCC 2040  
Qy 2068 GTCTGGCATAAATGATGATTTAGTGGTCACTAAAAAATTTCAAAGAAATTTCAAAGGCTAC 2127  
Db 2041 ATCTGGCATAAATGATGATTTAGTGGTCACTAAAAAATTTCAAAGAAATTTCAAAGGCTAC 2100  
Qy 2128 ATATCTGGAGCAGGAAAACTTCCACATCATTTGGAGGATCAGATCTAATAGCTCATCA 2187  
Db 2101 ATATCTGGAGCAGGAAAACTTCCACATCATTTGGAGGATCAGATCTAATAGCTCATCA 2160  
Qy 2188 TGCTCGAAGGATACAGAACTAGAGAGTGGCTAAGGCGAGGAATGGAGGTCAAAATCA 2247  
Db 2161 TGCTCGAAGGATACAGAACTAGAGAGTGGCTAAGGCGAGGAATGGAGGTCAAAATCA 2220



QY 2248 ACATGCAAAAGAGAGTCTCTTCTGATGATCTTTTTAGATACAATCCTTATTTAAABAG 2307  
DB |||||  
DB 2221 ACATGCAAAAGAGAGTCTCTTCTGATGATCTTTTTAGATACAATCCTTATTTAAABAG 2280  
QY 2308 GAGAAGATAACTGAGGATTTTAAAGAGAGCCATGGAAAACCTTCTAGTAGACATCTAC 2367  
DB |||||  
DB 2281 GAGAAGATAACTGAGGATTTTAAAGAGAGCCATGGAAAACCTTCTAGTAGACATCTAC 2340  
QY 2368 TTGAGGCCAACAGAGTTATATGAATATATAGTGTATAGAGCGATTTAAGTTACAATGTT 2427  
DB |||||  
DB 2341 TTGAGGCCAACAGAGTTATATGAATATATAGTGTATAGAGCGATTTAAGTTACAATGTT 2400  
QY 2428 TTATGCGCTAAATTTATTAATAAATGACAAAACTTTGATTTCTTTTGTATGTACAAT 2487  
DB |||||  
DB 2401 TTATGCGCTAAATTTATTAATAAATGACAAAACTTTGATTTCTTTTGTATGTACAAT 2460  
QY 2488 TGTGTTGTCGTGTTTTCAGGCTTTGTCATTTGTCATCTTTTTTTCATTTTTTAAATGTTTTG 2547  
DB |||||  
DB 2461 TGTGTTGTCGTGTTTTCAGGCTTTGTCATTTGTCATCTTTTTTTCATTTTTTAAATGTTTTG 2520  
QY 2548 TTTTATTAATAGTTAATATAGTCACAGTTCAAAATTTCTAAATTTCTAAGGTAAAGGAC 2607  
DB |||||  
DB 2521 TTTTATTAATAGTTAATATAGTCACAGTTCAAAATTTCTAAATTTCTAAGGTAAAGGAC 2579  
QY 2608 TAAAGTCACCTTCCACCATTTGCTAGCTACT- 2640  
DB |||||  
DB 2580 TAAAGTCACCTTCCACCATTTGCTAGCTACTTTGTTCCCTCAGAAAAAATTCATGA 2639  
QY 2641 -----TATTTTTA 2648  
DB |||||  
DB 2640 TACTCATTTCTTATGAATCTTTCCAGGATTTTGTAGTCCTATTCAAATTCCTATTTTTA 2699  
QY 2649 AATAATTTCTCACAAAATGATAGCATATAGCAGTGTCTACACCTGCTGTTTTTA 2708  
DB |||||  
DB 2700 AATAATTTCTCACAAAATGATAGCATATAGCAGTGTCTACACCTGCTGTTTTTA 2759  
QY 2709 CTTAGTAAGATTAATAATATAGGAATATCAATATAATGTTTTTAAATTTTTCTTTTC 2768  
DB |||||  
DB 2760 CTTAGT- AGATTAATAATATAGGAATATCAATATAATGTTTTTAAATTTTTCTTTTC 2818  
QY 2769 CATTTAGCTGTAGTCTTACCTAAACTCTGGTGATCCAAACAAAATGGCTTCAGTGTGCA 2828  
DB |||||  
DB 2819 CATTTAGCTGTAGTCTTACCTAAACTCTGGTGATCCAAACAAAATGGCTTCAGTGTGCA 2878  
QY 2829 GATGTCACTACATGTTATCTAGTACTAGAACTGAAGACATGTGGAGCTTCATCAA 2888  
DB |||||  
DB 2879 GATGTCACTACATGTTATCTAGTACTAGAACTGAAGACCATGTGGAGCTTCATCAA 2938  
QY 2889 ACATGGGTTTGTGTTTCCACAGAACTGGAAGACCTGTACCCTTTTGTGGTCTTTACTG 2948  
DB |||||  
DB 2939 ACATGGGTTTGTGTTTCCACAGAACTGGAAGACCTGTACCCTTTTGTGGTCTTTACTG 2998  
QY 2949 AGCTGGGTGGGTGCTCTGTTTGTAGCTTATTTAGAGTCTTAGGTTTCTCTACTTATAAGTA 3008  
DB |||||  
DB 2999 AGCTGGGTGGGTGCTCTGTTTGTAGCTTATTTAGAGTCTTAGGTTTCTCTACTTATAAGTA 3058  
QY 3009 GAAATGGTGAGATGTTGTTTCTTTTCTACCKTAAAGGAGATGGTAAGAAAACATGAATG 3068  
DB |||||  
DB 3059 GAAATGGTGAGATGTTGTTTCTTTTCTACCTTAAAGGAGATGGTAAGAAAACATGAATG 3118  
QY 3069 TCTTTTTCACACTTTATTTGACAAGTATTTTCAAGTCTGTGTTCAAAAATATATTCATG 3128  
DB |||||  
DB 3119 TCTTTTTCACACTTTATTTGACAAGTATTTTCAAGTCTGTGTTCAAAAATATATTCATG 3178  
QY 3129 TACCTGTGATCCAGCAAGGAGTTCCAGTCAAGAGTCACCTACCACTGATTTAGTTGTT 3188  
DB |||||  
DB 3179 TACCTGTGATCCAGCAAGGAGTTCCAGTCAAGAGTCACCTACCACTGATTTAGTTGTT 3238  
QY 3189 TAGAGAAATGGAATGGAACAGTGAGGAATGGAGGCCATATTTCCATGATCTTCCCTTGT 3248  
DB |||||  
DB 3239 TAGAGAAATGGAATGGAACAGTGAGGAATGGAGGCCATATTTCCATGATCTTCCCTTGT 3298  
QY 3249 AACAGAGCAACAGAGGGACAGAGGCTGGCCTCTACATCACTCTCACCTTCCAAATCT 3308

DB 3299 AACAGAGCAACAGAGGGACAGAGGCTGGCCTCTACATCACTCTCACCTTCCAAATCT 3358  
QY 3309 TGTGAAAGTGCACTACTCTTGCAGAACCAAAATTAACCTTACTTCCAAAGTTCTGGCTGCTTG 3368  
DB 3359 TGTGAAAGTGCACTACTCTTGCAGAACCAAAATTAACCTTACTTCCAAAGTTCTGGCTGCTTG 3418  
QY 3369 CAGGTGGAACTCCAGCTGCMAAGGAGTTAGGGAANTGAAGTCTTTTTTAAAGCTTCT 3428  
DB 3419 CAGGTGGAACTCCAGCTGCMAAGGAGTTAGGGAANTGAAGTCTTTTTTAAAGCTTCT 3478  
QY 3429 CAGCCTTCTTAGGGAACAGAAATTTGGGTGAGCCAATCTGCAATTTCTACTACAGCAATG 3488  
DB 3479 CAGCCTTCTTAGGGAACAGAAATTTGGGTGAGCCAATCTGCAATTTCTACTACAGCAATG 3538  
QY 3489 AGACAGTTAGATTAATTGAAATATTTATAGAGATTTATGAACACTTTAAATTTATGATAGTG 3548  
DB 3539 AGACAGTTAGATTAATTGAAATATTTATAGAGATTTATGAACACTTTAAATTTATGATAGTG 3598  
QY 3549 TATGACATTTGGATAGAACATGGGATCTTTAGNAGTAGAATTTGACAGGGCATATTAGTTG 3608  
DB 3599 TATGACATTTGGATAGAACATGGGATCTTTAGNAGTAGAATTTGACAGGGCATATTAGTTG 3658  
QY 3609 ATGAAATGGAGTCATTTGAGTCTTTTAAATAGCCATGTATCATAAATTCACCAAGTGAAGCTG 3668  
DB 3659 ATGAAATGGAGTCATTTGAGTCTCTTAAATAGCCATGTATCATAAATTCACCAAGTGAAGCTG 3718  
QY 3669 GTGGAACATATGCTCCTCAATTTTACAGTTAAGGAATATAAGGAGAGATTAATATGTTT 3728  
DB 3719 GTGGAACATATGCTCCTCAATTTTACAGTTAAGGAATATAAGGAGAGATTAATATGTTT 3778  
QY 3729 TCTGTCATGCCCCACAACTCCCTTCTTAAGGAGAGCTGCCCTACTATAGCAGTTTTTATAT 3788  
DB 3779 TCTGTCATGCCCCACAACTCCCTTCTTAAGGAGAGCTGCCCTACTATAGCAGTTTTTATAT 3838  
QY 3789 TGTCAATTTATGAATATAATGAATGAGGAGTTCTGGTACCTCCTCTCTTTTACAAATATG 3848  
DB 3839 TGTCAATTTATGAATATAATGAATGA- GAGTCTGCTGCTGCTCTCTTTTACAAATATG 3897  
QY 3849 GGTGTTGTCCAGTATTTTTCCCTTTTAAACCMWTCCTCAATTCGGGTGTGAGTGGATGT 3908  
DB 3898 G----TGTTGTGAGTATTTTCTTTTAAACCAATTCGGGTGTGAGTGA- ---- 3948  
QY 3909 TTCCAAATTTGGGTTTTAAATTTGTATATCCCTGATAGTATTAATGGGTCTATAGAAATCTT 3968  
DB 3949 TTTCAATTTGGTTTTAAATTTGTATATCCCTGATAGTATAATTTGGGTCTATAGAAATCTT 4008  
QY 3969 TATACATTTCTAGATCAAGTCTCTTTGCGGATATACGTATTTGAGATATTTACACCTAGTCT 4028  
DB 4009 TATACATTTCTAGATCAAGTCTCTTTGCGGATATATGATTTAGATATTTACACCTAGTCT 4068  
QY 4029 GTGGCTGACCTGTTTTCTTTATGCTTTTGTAGTAATAGAGTTTAAATTTTGAACAAGT 4088  
DB 4069 GTGGCTGACCTGTTTTCTTTATGCTTTTGTAGTAATAGAGTTTAAATTTTGAACAAGT 4128  
QY 4089 CAAATTTATTTTTCTTTGTTGATATTTTTCTCTCAATTTAAACCCCAAGATTTCA 4148  
DB 4129 CAAATTTATTTTTCTTTGTTGATATTTTTCTCTCAATTTAAACCCCAAGATTTCA 4188  
QY 4149 GATATTTCTGCTCTATATATATAAACTTTATATTTTTTATATTTGTGATCTACCTTGAATGA 4208  
DB 4189 GATATTTCTGCTCTATATATATAAACTTTATATTTTTTATATTTGTGATCTACCTTGAATGA 4248  
QY 4209 TATGATGTTGTGAATATAGGATCAGGTTCTTTTTTCCCCCATACAGATATCCAGTCA 4268  
DB 4249 TATGATGTTGTGAATATAGGATCAGGTTCTTTTTTCCCCCATACAGATATCCAGTCA 4308  
QY 4269 TTGTAACTGTTTATTTGAAGATTTATCTCTTCTCAATTTAAATTTACCTTGGCAATAGT 4328  
DB 4309 TTGTAACTGTTTATTTGAAGATTTATCTCTTCTCAATTTAAATTTACCTTGGCAATAGT 4368  
QY 4329 AAAAAATCAATTTAAACCAT 4346  
DB |||||



```
Db 4369 AAAAAATCAATTAAACCAT 4386
RESULT 4
AAZ89046
ID AAZ89046 standard; DNA; 65921 BP.
XX
AC AAZ89046;
XX
DT 01-JUN-2000 (first entry)
XX
DE Human nibrin DNA.
XX
KW Nibrin; human; DNA double strand break repair protein; diagnosis;
XX therapy; Nijmegen Breakage Syndrome; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN DE19818680-Cl.
XX
PD 09-MAR-2000.
XX
PF 27-APR-1998; 98DE-01018680.
XX
PR 27-APR-1998; 98DE-01018680.
XX
PA (UYBE ) UNIV BERLIN HUMBOLDT.
XX
DR WPI; 2000-196117/18.
XX
PT A DNA double strand break repair protein, Nibrin, and related DNA useful
PT for diagnosis and therapy of Nijmegen Breakage Syndrome and other
PT diseases influenced by DNA-double-strand break repair.
XX
PS Claim 4; Fig 1; 32pp; German.
XX
CC This invention describes a novel DNA double strand break repair protein,
CC Nibrin. Nibrin and DNA encoding it are useful for diagnosis and/or
CC therapy of diseases influenced by repair of DNA-double strand breaks, in
CC particular Nijmegen Breakage Syndrome. The product of the invention has
CC applications in gene therapy. This sequence encodes the nibrin protein
CC described in the invention
XX
SQ Sequence 65921 BP; 19501 A; 11699 C; 12463 G; 22149 T; 0 U; 109 Other;

Query Match 42.3%; Score 1862.8; DB 3; Length 65921;
Best Local Similarity 94.1%; Pred. No. 0;
Matches 2024; Conservative 9; Mismatches 26; Indels 91; Gaps 6;

QY 2278 TCTTTTATGATACCAATCTTATTTAAAGGAGAGAGATAACTGAGGATTTTAAAGAGAG 2337
DB 58191 TCTTTCAGATACCAATCTTATTTAAAGGAGAGAGATAACTGAGGATTTTAAAGAGAG 58250
QY 2338 CCATGAGAAATCTCTAGTAGCATCTACTTCCAGGCCCAACAGGTTATATGATATATA 2397
DB 58251 CCATGAGAAATCTCTAGTAGCATCTACTTCCAGGCCCAACAGGTTATATGATATATA 58310
QY 2398 GTCTATAGAGCCGATTTAAGTTACAATGTTTATGCGCTAAATTTTAAATTAATGCA 2457
DB 58311 GTCTATAGAGCCGATTTAAGTTACAATGTTTATGCGCTAAATTTTAAATTAATGCA 58370
QY 2458 CAAATCTTGTATCTTTTGTATGATCAATTTGTTGTCGTGTTTCCAGGCTTTGTCAATTG 2517
DB 58371 CAAATCTTGTATCTTTTGTATGATCAATTTGTTGTCGTGTTTCCAGGCTTTGTCAATTG 58430
QY 2518 CATCTTTTTCATTTTAAATGCTTTGTTTATTTAAATAGTATATAGTCACAGTTC 2577
DB 58431 CATCTTTTTCATTTTAAATGCTTTGTTTATTTAAATAGTATATAGTCACAGTTC 58490
QY 2578 AAAATCTAAATAGTAAAGGATAAGTCAACCTTCCACCATTTGCTTAGCT 2637
DB 58491 AAAATCTAAATAGTAAAGGATAAGTCAACCTTCCACCATTTGCTTAGCT 58549

QY 2638 ACT----- 2640
|||
DB 58550 ACTTGGTCCCTCAGAAAAAATTCATGATACTCATTTTATGAAATCTTTCCAGGAT 58609
QY 2641 -----TATTTTAAATAATTTCTACACAAATGATAGCATAAC 2678
|||
DB 58610 TTTTGAGTCTTATTCAAATTCCTATTTTAAATAATTTCTACACAAATGATAGCATAAC 58669
|||
QY 2679 ATATGCAGTGTCTACACCTTGCTTTTATCTAGTAAGATTAAAAATTTATAGGAATATC 2738
|||
DB 58670 ATATGCAGTGTCTACACCTTGCTTTTATCTAGT-AGATTAATAATTTATAGGAATATC 58728
|||
QY 2739 AATATAATGTTTAAATAATTTTCTTTCCATATATGCTGTAGTCTTACCTAACTCTGG 2798
|||
DB 58729 AATATAATGTTTAAATAATTTTCTTTCCATATATGCTGTAGTCTTACCTAACTCTGG 58788
|||
QY 2799 TGATCCAAACAAATGGCTTCAGTGTGAGATGTCACCTACATGTTATTTCTAGTACTAG 2858
|||
DB 58789 TGATCCAAACAAATGGCTTCAGTGTGAGATGTCACCTACATGTTATTTCTAGTACTAG 58848
|||
QY 2859 AAATCTGAAGACCATGTGAGACTTTCATCAAAACATGGTTTGTAGTTTTCACAGAAATGAAA 2918
|||
DB 58849 AAATCTGAAGACCATGTGAGACTTTCATCAAAACATGGTTTGTAGTTTTCACAGAAATGAAA 58908
|||
QY 2919 GACCTGTACCCCTTTTGGTGTCTTACGTAGCTGGTGGTGTCTGTTTGTAGCTTATTT 2978
|||
DB 58909 GACCTGTACCCCTTTTGGTGTCTTACGTAGCTGGTGGTGTCTGTTTGTAGCTTATTT 58968
|||
QY 2979 TAGAGTCTAGTTTCTCTACTTATTAAGTAGAATGCTGAGATTTCTTTCTTTCTTCTACC 3038
|||
DB 58969 TAGAGTCTAGTTTCTCTACTTATTAAGTAGAATGCTGAGATTTCTTTCTTTCTTCTACC 59028
|||
QY 3039 KTAAGGGAGAGATGTAAGAAAAACAATGAATGCTCTTTTTCAAAACCTTTATTTGACAAGTAT 3098
|||
DB 59029 TTAAGGGAGAGATGTAAGAAAAACAATGAATGCTCTTTTTCAAAACCTTTATTTGACAAGTAT 59088
|||
QY 3099 TTCAAGTCTGTGTCAAAAATATATTCATGTACCTGTGTATCCAGCAAGAGGAGTTCCA 3158
|||
DB 59089 TTCAAGTCTGTGTCAAAAATATATTCATGTACCTGTGTATCCAGCAAGAGGAGTTCCA 59148
|||
QY 3159 GTCAAGTCTGTGTCAAAAATATATTCATGTACCTGTGTATCCAGCAAGAGGAGTTCCA 3218
|||
DB 59149 GTCAAGTCTGTGTCAAAAATATATTCATGTACCTGTGTATCCAGCAAGAGGAGTTCCA 59208
|||
QY 3219 GGAGGCCATATTTCCATGACTTCCCTTGTAAACAGAGCAACAGAGGAGGAGGCTG 3278
|||
DB 59209 GGAGGCCATATTTCCATGACTTCCCTTGTAAACAGAGCAACAGAGGAGGAGGCTG 59268
|||
QY 3279 GCCTCTACATCACTCTCACCTTCCAAATCTTGTGGAAGTGCATCTACTTGCAGAACCAA 3338
|||
DB 59269 GCCTCTACATCACTCTCACCTTCCAAATCTTGTGGAAGTGCATCTACTTGCAGAACCAA 59328
|||
QY 3339 ATTAATCTTCTCCAGTCTGCTGCTGAGGTGGAATCTCCAGCTGCAAGGAGTTAG 3398
|||
DB 59329 ATTAATCTTCTCCAGTCTGCTGCTGAGGTGGAATCTCCAGCTGCAAGGAGTTAG 59388
|||
QY 3399 GGAATTAAGAGTCTTTTAAAGCTTCTCAGCCTTCCAGGCAACAGAAATTTGGGTCA 3458
|||
DB 59389 GGAATTAAGAGTCTTTTAAAGCTTCTCAGCCTTCCAGGCAACAGAAATTTGGGTCA 59448
|||
QY 3459 GCCAATCTGCAATTTCTACTACAGCATTTGAGACCATTTAGATTTTGAATATTTATAGA 3518
|||
DB 59449 GCCAATCTGCAATTTCTACTACAGCATTTGAGACCATTTAGATTTTGAATATTTATAGA 59508
|||
QY 3519 GAGTTATGACACTTAAATATATGATGTTAGTATGATTTGATAGACATGAGGATCTTT 3578
|||
DB 59509 GAGTTATGACACTTAAATATATGATGTTAGTATGATTTGATAGACATGAGGATCTTT 59568
|||
QY 3579 AGAAGTAGAATTTGACAGGCAATTTAGTTGATGAAATGAGCATTTTGAAGTCTTTTAAATA 3638
|||
DB 59569 AGAAGTAGAATTTGACAGGCAATTTAGTTGATGAAATGAGCATTTTGAAGTCTTTTAAATA 59628
|||
QY 3639 GCCATGTATCATTAATTAACCAAGTGAAGCTGGTGGAAACATATGCTCTCCATTTTACAGTTTA 3698
|||
```







```
Db 487 TTTAAATCAAGCTATATTGCAACTTGGAGGATTTACTGTAAACAATTTGGACAGAGGATG 546
Qy 493 CACTCACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 552
Db 547 CACTCACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 606
Qy 553 TGGACGTCCTCAATTTGTAAGCCAGAAATTTTAACTGTAATTTCTGAAAGCAGTTCAGTCCAA 612
Db 607 TGGACGTCCTCAATTTGTAAGCCAGAAATTTTAACTGTAATTTCTGAAAGCAGTTCAGTCCAA 666
Qy 613 GAAGCAGCTCCCAAAATTTGTAAGCCAGAAATTTTAACTGTAATTTCTGAAAGCAGTTCAGTCCAA 672
Db 667 GAAGCAGCTCCCAAAATTTGTAAGCCAGAAATTTTAACTGTAATTTCTGAAAGCAGTTCAGTCCAA 726
Qy 673 TAAAAATGTTGATCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 732
Db 727 TAAAAATGTTGATCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 786
Qy 733 TATATTTTGAATGCTCAAAACAGCATAGAAATTTGAGTTCGCGAGTGTCTTTTGGAGGTGG 792
Db 787 TATATTTTGAATGCTCAAAACAGCATAGAAATTTGAGTTCGCGAGTGTCTTTTGGAGGTGG 846
Qy 793 GGAAGCTAGCTTGTATACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 852
Db 847 GGAAGCTAGCTTGTATACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 906
Qy 853 AACGTGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 912
Db 907 AACGTGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 966
Qy 913 GAAATGGATTCAGTCAATATGATATGCTCCAAAGCAGGCTTTAGACCTATTCCTGA 972
Db 967 GAAATGGATTCAGTCAATATGATATGCTCCAAAGCAGGCTTTAGACCTATTCCTGA 1026
Qy 973 AGCAGAAATGGATTTGGCGTGTATTTTCATGACTACAAAGAAATTTACTGTGATCCTCAGGG 1032
Db 1027 AGCAGAAATGGATTTGGCGTGTATTTTCATGACTACAAAGAAATTTACTGTGATCCTCAGGG 1086
Qy 1033 CCATCCAGTACAGGATTTAAAGACAACAACCTCAGGACCAAGCTTTTCAAGAGCGGTGTC 1092
Db 1087 CCATCCAGTACAGGATTTAAAGACAACAACCTCAGGACCAAGCTTTTCAAGAGCGGTGTC 1146
Qy 1093 AGTTGATGAAAACTAATGCTCAAGCGCCCGAGTGAACACTTACACATACAGTACGTGACAC 1152
Db 1147 AGTTGATGAAAACTAATGCTCAAGCGCCCGAGTGAACACTTACACATACGTGACAC 1206
Qy 1153 AGAATCAGACGACGATACATGGGATTTGAGTGAAGGCCAAAGAAATCAAAGTCTC 1212
Db 1207 AGAATCAGACGACGATACATGGGATTTGAGTGAAGGCCAAAGAAATCAAAGTCTC 1266
Qy 1213 CAAAATGGAAACAAAAATTCAGAAATGCTTTTCAAGAGCGCACCCACTGTAAAGGAGTCTG 1272
Db 1267 CAAAATGGAAACAAAA----- 1282
Qy 1273 CAAAACAAGCTCTAATAATAATAGTATGATGATCAAAATFACTTTGGCTAAGATGAGAAATCCC 1332
Db 1283 ----- 1282
Qy 1333 AAATATACAGCTTTTACCACAACTAAATTTGCCAGTATATAATAAAGTAAAGATAGGCGTTC 1392
Db 1283 -----CACCACAACTAAATTTGCCAGTATATAATAAAGTAAAGATAGGCGTTC 1328
Qy 1393 TCAGCAGCAGCAGCAGCAACTCCATCAGAACTACTTTTCCAGCGCTCTACC-AAAAAAGGG 1451
Db 1329 TCAGCAGCAGCAGCAGCAACTCCATCAGAACTACTTTTCCAGCGCTCTACC-----AAGGG 1388
Qy 1452 AAAGGGATGAAGAAATCAAGAAATGCTTTTCATGCAAAATCAGCAGAAATAGAAACGCTCTT 1511
Db 1389 AAAGGGATGAAGAAATCAAGAAATGCTTTTCATGCAAAATCAGCAGAAATAGAAACGCTCTT 1448
Qy 1512 GTTCTCTTTTAGAACAAAAACAACCTGCTACACCCCTCATTTGTGGAAAAAATAAGGAGGAGC 1571
```

```
Db 1449 GTTCTCTTTTAGAACAAAAACAACCTGCTGTACACCCCTCATTTGTGGAAAAAATAAGGAGCAGC 1508
Qy 1572 ATCTATCTGAGATGAGCCTGTGGACACAACTCAGACAAATAACTTTATTATTACAGATACAG 1631
Db 1509 ATCTATCTGAGATGAGCCTGTGGACACAACTCAGACAAATAACTTTATTATTACAGATACAG 1568
Qy 1632 ATTTAAATCTATTGTGAAAAAATTTCTGCCAGTAAATCTCATGCTGCAGAAAAAGCTAAGAT 1591
Db 1569 ATTTAAATCTATTGTGAAAAAATTTCTGCCAGTAAATCTCATGCTGCAGAAAAAGCTAAGAT 1628
Qy 1692 CAAATAAAAAAGGGAAATGGGATGATGTGCCCATAGAGATGAAGTATTGGAACAGTTAT 1751
Db 1629 CAAATAAAAAAGGGAAATGGGATGATGTGCCCATAGAGATGAAGTATTGGAACAGTTAT 1688
Qy 1752 TCAAGCAGACAAAAACAGAGTTTAGAAATTTGATGTGAAAGTTCAAAAAACAGAGGAGAAAGATG 1811
Db 1689 TCAAGCAGACAAAAACAGAGTTTAGAAATTTGATGTGAAAGTTCAAAAAACAGAGGAGAAAGATG 1748
Qy 1812 TCAATTTTAGAAAAAGGCCCAAGGATGGATATAGAAAACAAATGACACTTTTCAGTGTGAAG 1871
Db 1749 TCAATTTTAGAAAAAGGCCCAAGGATGGATATAGAAAACAAATGACACTTTTCAGTGTGAAG 1808
Qy 1872 CAGTACCAGAAAGTAGCAAAATATCTCAAGAAAAATGAAATTTGGGAAAGAAACGTGAACCTCA 1931
Db 1809 CAGTACCAGAAAGTAGCAAAATATCTCAAGAAAAATGAAATTTGGGAAAGAAACGTGAACCTCA 1868
Qy 1932 AGGAAGACTCAGTATGCTCAGCTTAAAGAAATATCTAACTCAATGACAACTTCAGGATGATA 1991
Db 1869 AGGAAGACTCAGTATGCTCAGCTTAAAGAAATATCTAACTCAATGACAACTTCAGGATGATA 1928
Qy 1992 GTGAGATGCTTCCAAAAAGCTGTTATTGACTGAATTTAGATCAGTGTGATTTAAAAACT 2051
Db 1929 GTGAGATGCTTCCAAAAAGCTGTTATTGACTGAATTTAGATCAGTGTGATTTAAAAACT 1988
Qy 2052 CTACTTCCAGAAATCGTCTGCATAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2107
Db 1989 CTACTTCCAGAAATCGTCTGCATAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2044

RESULT 6
AAH03660
ID AAH03660 standard; cDNA; 752 BP.
XX
AC AAH03660;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:495.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
```











RESULT 9  
 AAI28794/c  
 ID AAI28794 standard; cDNA; 646 BP.  
 XX  
 AC AAI28794;  
 XX  
 DT 12-OCT-2001 (first entry)  
 XX  
 DE Colon tumour related determined cDNA sequence for clone 25908.  
 XX  
 KW Human; immunotherapy; diagnosis; colon cancer; colon tumour; immunogenic;  
 KW gene therapy; vaccine; colonic cancer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200149716-A2.  
 XX  
 PD 12-JUL-2001.  
 XX  
 PF 29-DEC-2000; 2000WO-US035596.  
 XX  
 PR 30-DEC-1999; 99US-00476296.  
 PR 10-JAN-2000; 2000US-00480321.  
 PR 15-FEB-2000; 2000US-00504629.  
 PR 06-MAR-2000; 2000US-00519444.  
 PR 19-MAY-2000; 2000US-00575251.  
 PR 29-JUN-2000; 2000US-00609448.  
 PR 28-AUG-2000; 2000US-00649811.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Xu J, Lodes MJ, Secretist H, Benson DR, Meagher MJ, Stolk JA;  
 PI King GE, Wang Y, Jiang Y;  
 XX  
 DR WPI; 2001-441847/47.  
 XX  
 PT Colon tumor associated proteins and nucleic acids useful for the  
 PT prevention, diagnosis and treatment of colonic cancer.  
 XX  
 PS Claim 25; Page 228; 472pp; English.  
 XX  
 CC The present invention describes colon tumour associated proteins (I) and  
 CC the polynucleotides (II) that encode them. (I) have cytostatic activity.  
 CC (I) and (II) can be used in gene therapy and vaccine production. (I) and  
 CC (II) may be used in the prevention, diagnosis and treatment of diseases  
 CC associated with inappropriate colon tumour associated protein (TCAP)  
 CC expression, such as colonic cancer. For example, (I) and (II) may be used  
 CC to treat disorders associated with decreased expression by rectifying  
 CC mutations or deletions in a patient's genome that affect the activity of  
 CC TCAPs by expressing inactive proteins or to supplement the patients own  
 CC production of them. Additionally, (II) may be used to produce the TCAP  
 CC proteins, by inserting the nucleic acids into a host cell culturing the  
 CC cell to express the protein. (II) and its complementary sequences may  
 CC also be used as DNA probes in diagnostic polymerase chain reaction (PCR)  
 CC and hybridisation assays to detect and quantitate the presence of similar  
 CC nucleic acids in samples, and therefore which patients may be in need of  
 CC restorative therapy. (I) may also be used as antigens in the production  
 CC of antibodies against TCAPs and in assays to identify modulators of TCAP  
 CC expression and activity. Anti-(I) antibodies and antagonists may also be  
 CC used to down regulate TCAP expression and activity. The anti-(I)  
 CC antibodies may also be used as diagnostic agents for detecting the  
 CC presence of TCAPs in samples (e.g. by enzyme linked immunosorbant assay  
 CC (ELISA)). AAI28460 to AAI29512 and AAM24494 to AAM24523 represent  
 CC nucleotide and amino acid sequences given in the exemplification of the  
 CC present invention  
 XX  
 SQ Sequence 646 BP; 239 A; 124 C; 95 G; 188 T; 0 U; 0 Other;  
 XX  
 Query Match 12.5%; Score 552.2; DB 4; Length 646;  
 Best Local Similarity 93.0%; Pred. No. 1.4e-106;  
 Matches 609; Conservative 4; Mismatches 32; Indels 10; Gaps 3;  
 QY 3419 AAAAGCTTCTCAGCCTTCCTAGGGAACAGAAATGGGTGAGCCAATCTGCAATTTCTACT 3478

Db ||||| 645 AAAGCTTCTCAGCCTTCCTAGGGAACAGAAATGGGAGGCCCAATCTGCAATTTCTACT 586  
 QY ||||| 3479 ACAGGCATTGAGACCAGTTAGATTATTTGAAATATTTATAGAGAGTTATGAACTTAATTT 3538  
 Db ||||| 585 ACAGGCATTGAGACCAGTTAGATTATTTGAAATATTTATAGAGAGTTATGAACTTAATTT 526  
 QY ||||| 3539 ATGATAGTGGTATGACATTGGATGAAACATATGGGATCTTTAGAACTAGAAATTTGACAGGC 3598  
 Db ||||| 525 ATGATAGTGGTATGACATTGGATGAAACATATGGGATCTTTAGAACTAGAAATTTGACAGGC 466  
 QY ||||| 3599 ATATTAGTTGATGAAATGGAGTCATTTCAGTCTCTTTAATAGCCATGTATCATTAATTTACCA 3658  
 Db ||||| 465 ATATTAGTTGATGAAATGGAGTCATTTCAGTCTCTTTAATAGCCATGTATCATTAATTTACCA 406  
 QY ||||| 3659 AGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGGAATATAATGGACAGATT 3718  
 Db ||||| 405 AGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGGAATATAATGGACAGATT 346  
 QY ||||| 3719 AATATTGTTTCTGTCATGCCACAAATCCCTTTCTAAGGAAGACTGCCCTACTATAGCAG 3778  
 Db ||||| 345 AATATTGTTTCTGTCATGCCACAAATCCCTTTCTAAGGAAGACTGCCCTACTATAGCAG 286  
 QY ||||| 3779 TTTTATATTTGTCAAATTTATGAATATGAATGAGGAGTTCTGGTACCTCTCTGCTTTT 3838  
 Db ||||| 285 TTTTATATTTGTCAAATTTATGAATATAATGAATGA-GAGTTCTGCTGCTCTCTGCTTTT 227  
 QY ||||| 3839 ACAAAATATGGTGGTGGTCCAGTATTTTCCCTTTTAAACMTTCCCAATTCGGGTGCTGT 3898  
 Db ||||| 226 ACAAAATATGG---TGTGTGTCAGTATTTTTCCTTTTAAACCAATTCCTGCTGTGTAG 171  
 QY ||||| 3899 AGTGGAGTGTTCCTCAATTTGGGTTTAAATTTGTATATCCCTGATAGCTATAATTTGGGTCA 3958  
 Db ||||| 170 TGATG-----TTTCAATTTGGTTTAAATTTGTATATCCCTGATAGCTATAATTTGGGTCA 116  
 QY ||||| 3959 AGAAATCTTTTATACATTTAGATGCAAGTCTCTTTGYCGGATPATACGTATTCAGATATTA 4018  
 Db ||||| 115 AGAAATCTTTTATACATTTAGATGCAAGTCTCTTTGTCGGATATATGTATTCAGATATTA 56  
 QY ||||| 4019 CACCTAGTCTGTGGCTTGACTGTTTCTTTATGCTCTTTTGATGAATAGAGTTTTT 4073  
 Db ||||| 55 CACCTAGTCTGTGGCTTGACTGTTTCTTTTATGCTCTTTTGATGAATAGAGTTTTT 1  
 RESULT 10  
 ABZ32980/c  
 ID ABZ32980 standard; cDNA; 646 BP.  
 AC ABZ32980;  
 XX  
 DT 30-JAN-2003 (first entry)  
 XX  
 DE Human colon tumour cDNA clone 25908 SEQ ID NO:343.  
 DE  
 KW Human; colon cancer; colon tumour; immunotherapy; diagnosis; cancer;  
 KW tumour; immune response; immunostimulant; cytostatic; vaccine; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200283070-A2.  
 XX  
 PD 24-OCT-2002.  
 XX  
 PF 09-APR-2002; 2002WO-US011475.  
 XX  
 PR 10-APR-2001; 2001US-00833263.  
 PR 03-AUG-2001; 2001US-00922217.  
 PR 19-DEC-2001; 2001US-00025380.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Xu J, Lodes MJ, Secretist H, Benson DR, Meagher MJ, Stolk JA;  
 PI Wang T, Jiang Y, Smith CL, King GE, Wang A, Clapper JD, Skeiky YAW;



PI	Fanger GR, Vedwick TS, Carter D;	
XX	WPI; 2003-067548/06.	
XX	New polynucleotide, useful for the preparation of a composition for	
PT	stimulating an immune response against, or treating, cancer.	
XX		
PS	Example 1; Page 231; 537pp; English.	
XX		
CC	The present invention describes compounds (I) for the immunotherapy and	
CC	diagnosis of colon cancer. Also described: (1) a method for detecting the	
CC	presence of cancer in a patient; (2) a method for stimulating and/or	
CC	expanding T cells specific for a tumour protein; (3) an isolated T cell	
CC	population comprising T cells prepared by the method of (2); (4) a method	
CC	for stimulating an immune response in a patient; (5) a method for	
CC	treating cancer in a patient; and (6) a method for inhibiting the	
CC	development of cancer in a patient. (I) have immunostimulant and	
CC	cytostatic activities and can be used in vaccines. AB232646 to AB233725	
CC	and ABP55343 to ABP55391 represent human colon cancer/tumour related	
CC	sequences used in the exemplification of the present invention	
XX		
SQ	Sequence 646 BP; 239 A; 124 C; 95 G; 188 T; 0 U; 0 Other;	
	Query Match 12.5%; Score 552.2; DB 8; Length 646;	
	Best Local Similarity 93.0%; Pred. No. 1.4e-106;	
	Matches 609; Conservative 4; Mismatches 32; Indels 10; Gaps 3;	
QY	3419 AAAAGCTTCTCAGCCTTCTTAGGGACAGAAATTTGGTGAGCCAAATCTGCAATTTCTACT	3478
DB		
QY	645 AAAGCTTCTCAGCCTTCTTAGGGACAGAAATTTGGTGAGCCAAATCTGCAATTTCTACT	586
DB		
QY	3479 ACAGGCATTGAGACCACTAGATTATTTGAATATTTATAGAGCTTATGACACTTAAAT	3538
DB		
QY	585 ACAGGCATTGAGACCACTAGATTATTTGAATATTTATAGAGCTTATGACACTTAAAT	536
DB		
QY	3539 ATGATAGTGTATGACATTTGGATAGAACATGGGATCTTTAGAGTAGAATTTGACAGGGC	3598
DB		
QY	525 ATGATAGTGTATGACATTTGGATAGAACATGGGATCTTTAGAGTAGAATTTGACAGGGC	466
DB		
QY	3599 ATATTAGTTCATGAATGGAGTCAATTTGAGTCTTTTAATAGCCATGTATCATTAATACCA	3658
DB		
QY	465 ATATTAGTTCATGAATGGAGTCAATTTGAGTCTTTTAATAGCCATGTATCATTAATACCA	406
DB		
QY	3659 AGTGAAGCTGGTGGAAACATATGGTCTCTCCATTTTACAGTTAAGGAATATATGACAGATT	3718
DB		
QY	405 AGTGAAGCTGGTGGAAACATATGGTCTCTCCATTTTACAGTTAAGGAATATATGACAGATT	346
DB		
QY	3719 AATATTGTTTCTGTCATGCCCAATCCCTTTCTAAGGAAGACTGCCCTTACTATAGCAG	3778
DB		
QY	345 AATATTGTTTCTGTCATGCCCAATCCCTTTCTAAGGAAGACTGCCCTTACTATAGCAG	286
DB		
QY	3779 TTTTATATTGTCATTTATGAATATATGAATGAGGAGTCTTGGTACCTCCTGCTTTT	3838
DB		
QY	285 TTTTATATTGTCATTTATGAATATATGAATGAGGAGTCTTGGTACCTCCTGCTTTT	227
DB		
QY	3839 ACAATATTGGGTGTTGTCAGATTTTTCCTTTTAAACMTTCCCAATTCGGGTGTG	3898
DB		
QY	226 ACAATATTGG-----TGTGTCAGATTTTTCCTTTTAAACATTCCAATCGGTGTG	171
DB		
QY	3899 AGGTGATGTTTCCATTTGGGTTTTAATTTGATATATCCCTGATAGCTATTAATTTGGGT	3958
DB		
QY	170 TGATG-----TTTCATTTGGTTTTAATTTGATATATCCCTGATAGCTATTAATTTGGGT	116
DB		
QY	3959 AGAATTTCTTATACATTTAGATGCAAGTCTCTTGGGATATAGCTATTTAGATTA	4018
DB		
QY	115 AGAATTTCTTATACATTTAGATGCAAGTCTCTTGGGATATAGCTATTTAGATTA	56
DB		
QY	4019 CACCTAGTCTGCGCTTGACTGTTTCTTTATGCTCTTTTATGATGAATAGAGTTTT	4073
DB		
QY	55 CACCTAGTCTGCGCTTGACTGTTTCTTTATGCTCTTTTATGATGAATAGAGTTTT	1
DB		



QY 722 GGGAAACATTATATTTTGAATGCCAACACAGCATAGAAATTGAGTTCCGAGTTGTC 781  
Db 321 GGGAAACATTATATTTTGAATGCCAACACAGCATAGAAATTGAGTTCCGAGTTGTC 262  
QY 782 TTTTGAGGTGGGAGCTAGGTTTGATAACAGAGAGAAATGAAGAGAAACATAATTTCTTT 841  
Db 261 TTTTGAGGTGGGAGCTAGGTTTGATAACAGAGAGAAATGAAGAGAAACATAATTTCTTT 202  
QY 842 TTGGCTCCGGGACGTTGTTGATACAGGAATAACAACTCACAGACCTTAATTCCT 901  
Db 201 TTGGCTCCGGGACGTTGTTGATACAGGAATAACAACTCACAGACCTTAATTCCT 142  
QY 902 GACTGTCAAGAAATGGATTTCAGTCAATAATGGATATCTCCAAAGGCAAGGCTTTAGA 961  
Db 141 GACTGTCAAGAAATGGATTTCAGTCAATAATGGATATCTCCAAAGGCAAGGCTTTAGA 82  
QY 962 CCTATTTCCTGAAGCAGAAATTTGGATTGGCGGTGATTTTCATGACTACAAAGAAATTAAGTCT 1021  
Db 81 CCTATTTCCTGAAGCAGAAATTTGGATTGGCGGTGATTTTCATGACTACAAAGAAATTAAGTCT 22  
QY 1022 GATCTTCAGGGCCATCCCAAGT 1042  
Db 21 GATCTTCAGGGCCATCCCAAGT 1  
RESULT 12  
ABL38148  
ID ABL38148 standard; cDNA; 544 BP.  
AC ABL38148;  
DT 08-APR-2002 (first entry)  
XX Human colon tumour antigen polynucleotide SEQ ID NO:1737.  
DE Human; colon cancer; colon tumour antigen; cytostatic; vaccine;  
KW colon tumour metastatic antigen; diagnosis; gene; ss.  
KW  
XX Homo sapiens.  
XX  
XX WO200196388-A2.  
XX  
XX 20-DEC-2001.  
XX  
XX 08-JUN-2001; 2001WO-US018557.  
XX  
XX 09-JUN-2000; 2000US-0210899P.  
PR 20-FEB-2001; 2001US-0270216P.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX  
PI Jiang Y, Harlocker SL, Secrist H;  
XX  
XX WPI; 2002-114514/15.  
DR  
XX  
PT Novel isolated colon tumor polynucleotide differentially expressed in  
PT colon tumor or colon metastatic tumor and polypeptides encoded by them,  
PT useful for inhibiting development of cancer in patient.  
XX  
XX  
PS Claim 1; SEQ ID NO 1737; 105pp; English.  
XX  
XX ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)  
CC which were isolated from human colon tumour and colon metastatic tumour  
CC cDNA libraries. (I) have cytoskeletal activity and can be used in vaccine  
CC production. (I) can be used for stimulating and/or expanding T cells  
CC specific for a tumour protein on contact with the T cells. They are also  
CC useful for inhibiting the development of cancer in a patient. (I) can be  
CC used as probes or primers for nucleic acid hybridisation, for preparing  
CC mutant species primers, or primers for use in genetic constructions. (I)  
CC can be used in the diagnosis of a colon tumour  
XX  
XX Sequence 544 BP; 173 A; 94 C; 125 G; 150 T; 0 U; 2 Other;

Query Match 12.2%; Score 538.6; DB 6; Length 544;  
Best Local Similarity 99.1%; Pred. No. 9.7e-104; Mismatches 2; Indels 0; Gaps 0;  
Matches 539; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 3130 ACCTGTGATCCAGCAAGAGGAGTTCCAGTCAAGAGTCACTACAACTGATTTAGTTGTTT 3189  
Db 1 ATCTGTGATCCAGCAAGAGGAGTTCCAGTCAAGAGTCACTACAACTGATTTAGTTGTTT 60  
QY 3190 AGAGAATGAGAAATGGAACAGTGGAGGAATGGAGGCCATATTTCCATGACTTCCCTTTGTA 3249  
Db 61 AGAGAATGAGAAATGGAACAGTGGAGGAATGGAGGCCATATTTCCATGACTTCCCTTTGTA 120  
QY 3250 ACAGAAGCAACAGAGAGGACAGAGGCTGAGGCTCTACATCACTCCCTTCCAAATCTT 3309  
Db 121 ACAGAAGCAACAGAGAGGACAGAGGCTGAGGCTCTACATCACTCCCTTCCAAATCTT 180  
QY 3310 GTGGAAGTGCATCTACTTTGCCAGAACCAAAATTAACCTTCTCCAAAGTTCTGGCTGCTGC 3369  
Db 181 GTGGAAGTGCATCTACTTTGCCAGAACCAAAATTAACCTTCTCCAAAGTTCTGGCTGCTGC 240  
QY 3370 AGGTGGAACTCCAGCTGCAAGGGAGTTAGGGAAATGAAGGTCCTTTTAAAAAGCTTCTC 3429  
Db 241 AGGTGGAACTCCAGCTGCAAGGGAGTTAGGGAAATGAAGGTCCTTTTAAAAAGCTTCTC 300  
QY 3430 AGCCTTCTCGGACAGAAATTTGGTGAGCAATCTGCNAATTTCTACTACAGGCATTGA 3489  
Db 301 AGCCTTCTCGGACAGAAATTTGGTGAGCAATCTGCNAATTTCTACTACAGGCATTGA 360  
QY 3490 GACCAGTTAGATTATTGAAATATTATAGAGAGTTATGAACACTTAAATTTATGATAGTGT 3549  
Db 361 GACCAGTTAGATTATTGAAATATTATAGAGAGTTATGAACACTTAAATTTATGATAGTGT 420  
QY 3550 ATGACATTTGGATAGAAACATGGGATCTTTAGAGTAGAATTGACAGGGCATTTAGTTGA 3609  
Db 421 ATGACATTTGGATAGAAACATGGGATCTTTAGAGTAGAATTGACAGGGCATTTAGTTGA 480  
QY 3610 TGAAATGGAGTCATTTGAGTCTTTTAATAGCCATGTCATATAATACCAGTGAAGCTGG 3669  
Db 481 TGAAATGGAGTCATTTGAGTCTTTTAATAGCCATGTCATATAATACCAGTGAAGCTGG 540  
QY 3670 TGGG 3673  
Db 541 TDNA 544  
RESULT 13  
AAH09188/c  
ID AAH09188 standard; cDNA; 543 BP.  
XX  
XX AC AAH09188;  
XX  
XX DT 26-JUN-2001 (first entry)  
XX  
XX Human cDNA clone (3'-primer) SEQ ID NO:6023.  
XX  
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
XX  
XX OS Homo sapiens.  
XX  
XX PN EP1074617-A2.  
XX  
XX PD 07-FEB-2001.  
XX  
XX 28-JUL-2000; 2000EP-00116126.  
PF  
XX 29-JUL-1999; 99JP-00248036.  
PR 27-AUG-1999; 99JP-00300253.  
PR 11-JAN-2000; 2000JP-00118776.  
PR 02-MAY-2000; 2000JP-00183767.  
PR 09-JUN-2000; 2000JP-00241899.  
XX  
XX (HELI-) HELIX RES INST.



XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX WPI; 2001-318749/34.  
XX  
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.  
PT  
PT  
XX  
XX  
PS Claim 3; SEQ ID NO 6023; 2537bp + Sequence Listing; English.  
XX  
CC The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises:  
CC (a) an oligo-dT primer and an oligonucleotide complementary to the  
CC complementary strand of a polynucleotide which comprises one of the 5602  
CC nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in the  
CC specification. The primer sets can be used in antisense therapy and in  
CC gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893  
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent  
CC oligonucleotides, all of which are used in the exemplification of the  
CC present invention  
XX  
SQ Sequence 543 BP; 128 A; 111 C; 83 G; 219 T; 0 U; 2 Other;

Query Match 11.8%; Score 520.4; DB 4; Length 543;  
Best Local Similarity 98.3%; Pred. No. 6.8e-100;  
Matches 535; Conservative 0; Mismatches 8; Indels 1; Gaps 1;  
1564 GGAGCAGCATCTATCTGAGAATGAGCGCTGGGACACAAACTCAGACAAATTAATTATTC 1623  
543 GNAGCAGCTTCTATCTGAGAATGAGCGCTG-GGACCCAAACTCAGCCAAATTAATTTC 485  
1624 AGTACAGATTTAAATCTATCTGAAAATTCGCCAGTAATCTCATGCTCCAGAAA 1683  
484 AGATCCAGATTTTAAATCTATCTGAAAATTCGCCAGTAATCTCATGCTCCAGAAA 425  
1684 GCTAAGATCAATATATAAAGGAAATGGATGATGCGCCATAGAGAGTGAAGTATTGGA 1743  
424 GCTAAGATCAATATATAAAGGAAATGGATGATGCGCCATAGAGAGTGAAGTATTGGA 365  
1744 ACAGTTATTCAAGGACACAAAACACAGATTAGAAATTTGATGTAAGATTTCAAAAACAGGA 1803  
364 ACAGTTATTCAAGGACACAAAACACAGATTAGAAATTTGATGTAAGATTTCAAAAACAGGA 305  
1804 GGNAGATGTCATGTTAGAAAAGGCCAAGATGGATATAGAAACAATGACACTTTCAG 1863  
304 GGAAGATGTCATGTTAGAAAAGGCCAAGATGGATATAGAAACAATGACACTTTCAG 245  
1864 TGATGAAGCAGTACCAAGAGTAGCAAAATATCTCAAGAAAATGAAATTTGGGAAGAACG 1923  
244 TGATGAAGCAGTACCAAGAGTAGCAAAATATCTCAAGAAAATGAAATTTGGGAAGAACG 185  
1924 TGAACCTCAAGGAGAGCTCACTATGTCAGCTAAGAAAATATCTAACAATGACAACTTCA 1983  
184 TGAACCTCAAGGAGAGCTCACTATGTCAGCTAAGAAAATATCTAACAATGCGCAACTTCA 125  
1984 GGATGATAGTAGAGATGCTTCCAAAAAGCTGTTATTGACTGAATTTAGATCACTGGTAT 2043  
124 GGATGATAGTAGAGATGCTTCCAAAAAGCTGTTATTGACTGAATTTAGATCACTGGTAT 65

QY 2044 TAAAAACTCTACTTCAGAAATCCGCTCGCATAAATGATGATTATGGTCAACTAAAAA 2103  
Db 64 TAAAAACTCTACTTCAGAAATCCGCTCGCATAAATGATGATTATGGTCAACTAAAAA 5  
QY 2104 TTTC 2107  
Db 4 TTTC 1  
RESULT 14  
ADST71935  
ID ADS71935 standard; cDNA; 587 BP.  
XX  
AC ADS71935;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Human kidney tumour specific cDNA, SEQ ID 532.  
XX  
KW Human; ss; kidney tumour; kidney cancer; cytostatic; gene therapy;  
KW T-cell; immune response.  
XX  
OS Homo sapiens.  
XX  
PN US2003109434-A1.  
PD 12-JUN-2003.  
XX  
PF 19-MAR-2002; 2002US-00102524.  
XX  
PR 19-MAR-2001; 2001US-0277245P.  
PR 21-DEC-2001; 2001US-0343340P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Algate PA, Mannion J, Gaiger A, Gordon B, Harlocker SL;  
XX  
DR WPI; 2002-759855/82.  
XX  
PT New isolated polynucleotides and polypeptides, useful for detecting the  
PT presence of, and treating cancer, particularly kidney cancer by  
PT stimulating T-cells specific for a tumor protein, and stimulating immune  
PT response in a patient.  
XX  
PS Claim 1; SEQ ID NO 532; 78bp; English.  
XX  
CC The invention relates to a new isolated polynucleotide (a Human kidney  
CC tumour specific cDNA) comprising any one of the 1855 sequences identified  
CC in the specification (or their complements, degenerate variants,  
CC sequences consisting of at least 20 contiguous residues them, sequences  
CC that hybridise to them under highly stringent conditions or sequences  
CC having at least 75 or 90% sequence identity to the 1855 sequences. Also  
CC included are detecting/determining the presence of cancer in a patient,  
CC stimulating an immune response in a patient; treating kidney cancer in a  
CC patient, an isolated polypeptide encoded by one of the 1855 sequences, an  
CC expression vector comprising the polynucleotide operably linked to an  
CC expression control sequence, a host cell transformed/transfected with the  
CC vector, an isolated antibody (or its antigen-binding fragment) that  
CC specifically binds to the protein, a fusion protein comprising at least  
CC one of the proteins, stimulating and/or expanding T-cells specific for a  
CC tumour protein, an isolated T-cell population comprising the T-cells, a  
CC composition comprising a first component (such as a carrier or  
CC immunostimulant) and a second component (comprising one of the  
CC polynucleotides, the polypeptides, an antibody, T-cell or an antigen-  
CC presenting cell that expresses the polynucleotide) and a diagnostic kit  
CC comprising at least one of the oligonucleotides, or at least one antibody  
CC and a detection reagent comprising a reporter group. The polynucleotides,  
CC polypeptides, antibodies and antigen-presenting cells are useful for  
CC detecting the presence of, and treating cancer, particularly kidney  
CC cancer by stimulating and/or expanding T-cells specific for a tumour  
CC protein, and stimulating immune response in a patient. The present  
CC sequence is one of the Human kidney tumour specific cDNAs. Note: The



CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from USPTO  
CC at seqdata.uspto.gov/sequence.html?DocID=20030109434.

XX SQ Sequence 587 BP; 177 A; 86 C; 114 G; 210 T; 0 U; 0 Other;

Query Match 11.5%; Score 507; DB 7; Length 587;  
Best Local Similarity 93.6%; Pred. No. 4.7e-97;  
Matches 559; Conservative 4; Mismatches 24; Indels 10; Gaps 3;

QY 3419 AAAAGCTTCTCAGCCTTCCTAGGAAACAGAAATGGGTGAGCCAAATCTGCAATTTCTACT 3478

Db 1 AAAAGCTTCTCAGCCTTCTAGGAAACAGAAATGGGTGAGCCAAATCTGCAATTTCTACT 60

QY 3479 ACAGGCATTGAGACCAAGTTAGATTATTTGAAATATATAGAGAGTTATGAACACTTAAAT 3538

Db 61 ACAGGCATTGAGACCAAGTTAGATTATTTGAAATATATAGAGAGTTATGAACACTTAAAT 120

QY 3539 ATGATAGTGTATGACATTGGATAGAACATGGGATACCTTTAGAAGTAGAATTGACAGGGC 3598

Db 121 ATGATAGTGTATGACATTGGATAGAACATGGGATACCTTTAGAAGTAGAATTGACAGGGC 180

QY 3599 ATATTAGTTGATGAAATGAGTCAATTTGAGTCTCTTTTAATAGCCATGTATCATAATTACCA 3658

Db 181 ATATTAGTTGATGAAATGAGTCAATTTGAGTCTCTTTTAATAGCCATGTATCATAATTACCA 240

QY 3659 AGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGGAATATATGACACAGATT 3718

Db 241 AGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGGAATATATGACACAGATT 300

QY 3719 AATATTGTTCTGTCATGCCCAATCCCTTCTAAGGAAGACTGCCCTACTATAGCAG 3778

Db 301 AATATTGTTCTGTCATGCCCAATCCCTTCTAAGGAAGACTGCCCTACTATAGCAG 360

QY 3779 TTTTATATTGTCAAATTTATGAATATGAATGAGGAGTTCTCGTACTCTCTGCTTT 3838

Db 361 TTTTATATTGTCAAATTTATGAATATGAATGAGTCTCGTACTCTCTGCTTT 419

QY 3839 ACAATATTGGGTGTTGTCAGATATTTTCCCTTTTAAACMTTCCCAATTCGGGTGT 3898

Db 420 ACAATATTGG-----TGTTGTCAGATATTTTCCCTTTTAAACCATTCCAATTCGGGTGTAG 475

QY 3899 AGTGGATGTTTCCATTTGGGTTTTAAATTTGTATATCCCTGTAGCTATAATTTGGGTGAT 3958

Db 476 TGAAG-----TTTCATTTTGGTTTTAAATTTGTATATCCCTGTAGCTATAATTTGGGTGAT 530

QY 3959 AGAAATCTTTATACATCTAGATGCAAGTCTCTTGYCGGATATACGTATTGAGATA 4015

Db 531 AGAAATCTTTATACATCTAGATGCAAGTCTCTTGYCGGATATATGATTGAGATA 587

RESULT 15

ID ACH26622

XX ACH26622 standard; cDNA; 468 BP.

AC ACH26622;

XX 13-OCT-2003 (first entry)

DT Human adult ovary cDNA #5002.

XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;

XX Genome mapping; biodiversity; genetic disorder.

XX Homo sapiens.

XX US2003073623-A1.

XX 17-APR-2003.

XX 30-JUL-2001; 2001US-00918995.

XX 30-JUL-2001; 2001US-00918995.

XX

PA (DRMA/) DRMANAC R T.

PA (LABA/) LABAT I.

PA (STAC/) STACHE-CRAIN B.

PA (DICK/) DICKSON M C.

PA (JONE/) JONES L W.

XX

PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

XX

WPI; 2003-615964/58.

XX

New polynucleotide sequences obtained from various cDNA libraries, useful as hybridization probes, as oligomers for PCR, for chromosome and gene mapping, in the recombinant production of protein, or in generating antisense DNA or RNA.

XX

Claim 1; SEQ ID NO 13834; 44pp; English.

XX

The invention relates to an isolated polynucleotide comprising any one of 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was determined by the technique of SBH (sequencing by hybridisation). Also included is a purified polypeptide comprising a sequence corresponding to a reading frame of the novel polynucleotide. The nucleic acid sequences are useful in diagnostics as expressed sequence tags (EST) for identifying expressed genes or for physical mapping of the human genome, in forensics, in assessing biodiversity, or in identifying mutations responsible for genetic disorders and other traits. The nucleotide sequences are also useful as hybridisation probes, as oligomers for PCR, for chromosome and gene mapping, in the recombinant production of protein, or in generating antisense DNA or RNA. The purified polypeptide is useful for generating antibodies specific for it. The present sequence is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030073623

XX

Seqdata.uspto.gov/sequence.html?DocID=20030073623

XX

Query Match 9.2%; Score 405.6; DB 9; Length 468;

Best Local Similarity 97.4%; Pred. No. 1.2e-75;

Matches 411; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY

2938 TGGTCTTACTGAGCTGGGTGGTGTCTGTTTGGAGCTATTATAGAGTCTCTAGTTTTCCTA 2997

Db

47 TGGTCTTACTGAGCTGGGTGGTGTCTGTTTGGAGCTATTATAGAGTCTCTAGTTTTCCTA 106

QY

2998 CTTATAAAGTAGAAAATGGTGAGATTGTTTTCTTTTCTACCKTAAAGGAGATGGTAAGA 3057

Db

107 CTTATAAAGTAGAAAACGGGGAGATTGTTTTGTTTTTCTACCTTAAAGGAGATGGCAAGA 166

QY

3058 AACAAATGAATGTCCTTTTTTCAAACTTTATGACAAGTGATTTTCAAGTCTGTGTTCAAAA 3117

Db

167 AACAAATGAATGTCCTTTTTTCAAACTTTATGACAAGTGATTTTCAAGTCTGTGTTCAAAA 226

QY

3118 ATATATTTCATGTACCTGTGATCCAGCAAGAGGGAGTTCCAGTCAAGAGTCACTCAACT 3177

Db

227 ATATATTTCATGTACCTGTGATCCAGCAAGAGGGAGTTCCAGTCAAGAGTCACTCAACT 286

QY

3178 GATTAGTTGTTTATAGAGAATGAGAAAATGGAACAGTGAAGGAATGGAGCCCATATTTCATGA 3237

Db

287 GATTAGTTGTTTATAGAGAATGAGAAAATGGAACAGTGAAGGAATGGAGCCCATATTTCATGA 346

QY

3238 CTTCCCTTGTAAACAGAACAGAGGGAACAGAGGCTGGCCTCTACATCACTCTCAC 3297

CC

CC

CC

XX

SQ

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

3;

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db



Search completed: August 30, 2005, 06:01:37  
Job time : 2165 secs



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2005, 05:25:38 ; Search time 677 Seconds  
(without alignments)  
10641.841 Million cell updates/sec

Title: US-09-837-138-1  
Perfect score: 4403  
Sequence: 1 tccggcacaggcgggttg.....accgggtgagctccagct 4403

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/prodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/prodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/prodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/prodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/prodata/1/ina/PCUTUS COMB.seq.\*  
6: /cgn2\_6/prodata/1/ina/backfileseq.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4108.4	93.3	4386	3	US-09-300-008B-1
2	552.2	12.5	646	4	US-09-401-064-343
3	84	1.9	601	4	US-09-949-016-79057
4	82.8	1.9	784019	4	US-09-949-016-14033
5	82.8	1.9	828152	4	US-09-949-016-12777
6	81.6	1.9	3285	4	US-09-573-080A-425
7	79.8	1.8	7218	1	US-08-232-463-14
8	78	1.8	75216	4	US-09-949-016-14061
9	75.4	1.7	141454	4	US-09-949-016-12055
10	75.4	1.7	141455	4	US-09-949-016-17190
11	74.8	1.7	173992	4	US-09-949-016-13379
12	69.6	1.6	612	4	US-09-902-540-1357
13	68.2	1.5	69709	4	US-09-949-016-15784
14	68.2	1.5	146039	4	US-09-949-016-12449
15	67.8	1.5	601	4	US-09-949-016-13149
16	67.8	1.5	601	4	US-09-949-016-143186
17	67.8	1.5	1141	4	US-09-806-708B-22
18	67.4	1.5	29357	4	US-09-949-016-16676
19	67.2	1.5	601	4	US-09-949-016-152075
20	67.2	1.5	601	4	US-09-949-016-152148
21	67.2	1.5	601	4	US-09-949-016-152221
22	67.2	1.5	601	4	US-09-949-016-152294
23	67.2	1.5	601	4	US-09-949-016-159380
24	67.2	1.5	601	4	US-09-949-016-159453
25	67.2	1.5	601	4	US-09-949-016-159526
26	67.2	1.5	601	4	US-09-949-016-159599
27	67.2	1.5	78846	4	US-09-949-016-12396

ALIGNMENTS

RESULT 1

US-09-300-008B-1

; Sequence 1, Application US/09300008B

; Patent No. 6458534

; GENERAL INFORMATION:

; APPLICANT: Concannon et al.

; TITLE OF INVENTION: A GENE ASSOCIATED WITH NIEMEN BREAKAGE

; FILE REFERENCE: SYNDROME, ITS GENE PRODUCT AND METHODS FOR THEIR USE

; CURRENT APPLICATION NUMBER: US/09/300,008B

; CURRENT FILING DATE: 1999-04-27

; PRIOR APPLICATION NUMBER: US 60/083,269

; PRIOR FILING DATE: 1998-04-27

; NUMBER OF SEQ ID NOS: 64

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1

; TYPE: DNA

; ORGANISM: Homo sapien

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (26)...(2287)

US-09-300-008B-1

Query Match 93.3%; Score 4108.4; DB 3; Length 4386;

Best Local Similarity 97.1%; Pred. No. 0;

Matches 4272; Conservative 7; Mismatches 28; Indels 91; Gaps 6;

QY	28	GCCCCAGCCCTCAGGAGCGGACCGATGTGGAACTGCTGCCCGCGCGGCGCGCGCGCGG	87	Sequence 12791, A
Db	1	GCCCCAGCCCTCAGGAGCGGACCGATGTGGAACTGCTGCCCGCGCGGCGCGCGCGG	60	Sequence 12792, A
QY	88	AGGAGAACCATACAGACTTTTGACTGGCGTTGAGTACGTTGTGGAGGAAAAAATGTGC	147	Sequence 12793, A
Db	61	AGGAGAACCATACAGACTTTTGACTGGCGTTGAGTACGTTGTGGAGGAAAAAATGTGC	120	Sequence 16013, A
QY	148	CATTCTAATTGAAATGATCAGTCGATCAGCCGAATCATGCTGTGTTAACTGCTAACTT	207	Sequence 16014, A
Db	121	CATTCTGATTGAAATGATCAGTCGATCAGCCGAATCATGCTGTGTTAACTGCTAACTT	180	Sequence 16015, A
QY	208	TTCTGTAAACCACTGAGTCAACAGATGAATCCCTGTATTGACATTAAGAATATTC	267	Sequence 16016, A
Db	181	TTCTGTAAACCACTGAGTCAACAGATGAATCCCTGTATTGACATTAAGAATATTC	240	Sequence 16201, A
QY	268	TAAGTATGTTACTCTTTGTTAATGAGGAAAAAATGAGAAATGCGATGCTTTTCCGAACTTGA	327	Sequence 16202, A
Db	241	TAAGTATGTTACTCTTTGTTAATGAGGAAAAAATGAGAAATGCGATGCTTTTCCGAACTTGA	300	Sequence 16203, A
QY	328	GTCCGGGGAGTGTATTACTTTTGGAGTGTTCGAGTAAATTCAGAAATAGATATGAGCC	387	Sequence 16204, A



Db 301 GTCCGGGATGGTATTACTTTTGGAGTGTGTTGGAAGTAAATTCAGAATAGAGTATGAGCC 360  
Qy 388 TTTGGTTGATGCTCTTCTTTTGTAGATGTCCTGGGAAAACTGCTTTAAATCAAGCTAT 447  
Db 361 TTTGGTTGATGCTCTTCTTTTGTAGATGTCCTGGGAAAACTGCTTTAAATCAAGCTAT 420  
Qy 448 ATTGCAACTTGGAGGATTTACTGTAAACAATTGGACAGGAAGATGCACTCACCTTGTCAT 507  
Db 421 ATTGCAACTTGGAGGATTTACTGTAAACAATTGGACAGGAAGATGCACTCACCTTGTCAT 480  
Qy 508 GGTATCAGTGAAGTTTACCATTAATAACAATATGTGCATCATTTGTGGACGTCCAATGT 567  
Db 481 GGTATCAGTGAAGTTTACCATTAATAACAATATGTGCATCATTTGTGGACGTCCAATGT 540  
Qy 568 AAGCCAGATATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAAGACGAGCTCCACA 627  
Db 541 AAGCCAGATATTTTACTGAAATTCCTGAAAGCAGTTTGAGTCCAAGACGAGCTCCACA 600  
Qy 628 AATTGAAAGTTTTCACCACTCTTGATGAACCATCTATTGGAAGTAAAAATGTTGATCT 687  
Db 601 AATTGAAAGTTTTCACCACTCTTGATGAACCATCTATTGGAAGTAAAAATGTTGATCT 660  
Qy 688 GTCAGGACGCGAGGAAGAAAAAATAATCTTCAAAGGGAATAATTTATATTTTGAATGC 747  
Db 661 GTCAGGACGCGAGGAAGAAAAAATAATCTTCAAAGGGAATAATTTATATTTTGAATGC 720  
Qy 748 CAACAGCATAGAAATTTGATTCGCGAGTTGTCCTTTGGAGTGGGGAAGCTAGTTGAT 807  
Db 721 CAACAGCATAGAAATTTGATTTCCGACGTTGCTTTGGAGTGGGGAAGCTAGTTGAT 780  
Qy 808 AACAGAAGAGATGAAGAAGAACATAATTTCTTTTGGCTCCGGGAACGTGTGTTGCA 867  
Db 781 AACAGAAGAGATGAAGAAGAACATAATTTCTTTTGGCTCCGGGAACGTGTGTTGCA 840  
Qy 868 TACAGGAATAACAACTCAAGACCTTAATTCCTGATGTCAAGAGAAATGGAATCAGTC 927  
Db 841 TACAGGAATAACAACTCAAGACCTTAATTCCTGATGTCAAGAGAAATGGAATCAGTC 900  
Qy 928 AATAATGGATATGCTCCAAAGCAAGGTCTTAGACCTATTCTGAAGCAGAAATTTGGATT 987  
Db 901 AATAATGGATATGCTCCAAAGCGAGGTCTTAGACCTATTCTGAAGCAGAAATTTGGATT 960  
Qy 988 GGGGTGATTTTCATGACTACAAAGAAATTAATCTGATCTCAGGGCCATCCCAAGTACAGG 1047  
Db 961 GGGGTGATTTTCATGACTACAAAGAAATTAATCTGATCTCAGGGCCATCCCAAGTACAGG 1020  
Qy 1048 ATTAAGACAACTCAAGGACCAAGCCCTTCAAGGCGTGTGATGATGAATAACT 1107  
Db 1021 ATTAAGACAACTCAAGGACCAAGCCCTTCAAGGCGTGTGATGATGAATAACT 1080  
Qy 1108 AATGCCAAGGCCCGAGTGAACACTCAACATACGTAGTGCACAGAAATCAGAGCAAGC 1167  
Db 1081 AATGCCAAGGCCCGAGTGAACACTCAACATACGTAGTGCACAGAAATCAGAGCAAGC 1140  
Qy 1168 AGATACATGGATTTGAGTGAAGGCAAGAAATCAAAGTCTCCAAAATGAACAAAA 1227  
Db 1141 AGATACATGGATTTGAGTGAAGGCAAGAAATCAAAGTCTCCAAAATGAACAAAA 1200  
Qy 1228 ATTCAAGATGCTTTCACAGACGACCCACTGCTGAAAGGAGTCTCGAAAACAGCTCTAA 1287  
Db 1201 ATTCAAGATGCTTTCACAGACGACCCACTGCTGAAAGGAGTCTCGAAAACAGCTCTAA 1260  
Qy 1288 TAATAATAGTATGGTATCAATACTTTGGTAAAGTGAAGATCCCAAACTATCAGCTTTC 1347  
Db 1261 TAATAATAGTATGGTATCAATACTTTGGTAAAGTGAAGATCCCAAACTATCAGCTTTC 1320  
Qy 1348 ACCAACTAAATTCGCAAGTATAATAAAGTAAAGATAGGGCTTCTCAGCAGCAGCAGAC 1407  
Db 1321 ACCAACTAAATTCGCAAGTATAATAAAGTAAAGATAGGGCTTCTCAGCAGCAGCAGAC 1380  
Qy 1408 CAACTCCATCAGAAACTACTTTCAGCCGCTACCAAAAAAAGGGAAGGATGAAGAAA 1467  
Db 1381 CAACTCCATCAGAAACTACTTTCAGCCGCTACCAAAAAAAGGGAAGGATGAAGAAA 1440

Qy 1468 TCAAGAAATGCTCTTCATCAATCAGCAAGATAGAAAGCTCTGTTCTCTTTTGAACA 1527  
Db 1441 TCAAGAAATGCTCTTCATCAATCAGCAAGATAGAAAGCTCTGTTCTCTTTTGAACA 1500  
Qy 1528 AACCAAACTGCTACACCCCTCATTTGTGAAATAAAGGAGCAGCATCTATCTGAGAATGA 1587  
Db 1501 AACCAAACTGCTACACCCCTCATTTGTGAAATAAAGGAGCAGCATCTATCTGAGAATGA 1560  
Qy 1588 GCCTGTGACACAAAATCAGACAATACTTATTTACAGATACAGATTTAAATCTATTGT 1647  
Db 1561 GCCTGTGACACAAAATCAGACAATACTTATTTACAGATACAGATTTAAATCTATTGT 1620  
Qy 1648 GAAAAATTTCTGCCAGTAAATCTCATGTCAGAAAAGCTTAAGATCAATATAAAAAAGGA 1707  
Db 1621 GAAAAATTTCTGCCAGTAAATCTCATGTCGAGAAAAGCTTAAGATCAATATAAAAAAGGA 1680  
Qy 1708 AATGATGATGTGGCCATAGAGATGAAGTATTGGAACTATTATTCAAGGACACAAAACC 1767  
Db 1681 AATGATGATGTGGCCATAGAGATGAAGTATTGGAACTATTATTCAAGGACACAAAACC 1740  
Qy 1768 AGAGTTAGAAATTTGATGTGAAAGTTCAAAAAACAGGAGGAAGATGTCATTTAGAAAAAG 1827  
Db 1741 AGAGTTAGAAATTTGATGTGAAAGTTCAAAAAACAGGAGGAAGATGTCATTTAGAAAAAG 1800  
Qy 1828 GCCAAGGATGGATATAGAAAACAAATGACACTTTTCAGTGATGAAGCAGTACCAGAAAGTAG 1887  
Db 1801 GCCAAGGATGGATATAGAAAACAAATGACACTTTTCAGTGATGAAGCAGTACCAGAAAGTAG 1860  
Qy 1888 CAAAAATCTCAAGAAAAATGAAATTTGGGAAGAAACGTGAACCTCAAGGAAGACTCATATG 1947  
Db 1861 CAAAAATCTCAAGAAAAATGAAATTTGGGAAGAAACGTGAACCTCAAGGAAGACTCATATG 1920  
Qy 1948 GTCAGCTAAAGAAATATCTAACATGACAAACTTCAGGATGATGATGAGATGCTTCCAAA 2007  
Db 1921 GTCAGCTAAAGAAATATCTAACATGACAAACTTCAGGATGATGATGAGATGCTTCCAAA 1980  
Qy 2008 AAGCTGTTTATTGACTGAAATTTAGATCACTGGTGATTAATAAACTCTATTCCAGAAATCC 2067  
Db 1981 AAGCTGTTTATTGACTGAAATTTAGATCACTGGTGATTAATAAACTCTATTCCAGAAATCC 2040  
Qy 2068 GTCTGGCATAAATGATGATTTAGTGTCAACTAAAAAATTTCAAGAAATTTCAAAAAGGTCA 2127  
Db 2041 ATCTGGCATAAATGATGATTTAGTGTCAACTAAAAAATTTCAAGAAATTTCAAAAAGGTCA 2100  
Qy 2128 ATATCTGGACAGCAAAAACCTTCCACATCATTTGGAGGATCAGATCTAATAGTCTATCA 2187  
Db 2101 ATATCTGGACAGCAAAAACCTTCCACATCATTTGGAGGATCAGATCTAATAGTCTATCA 2160  
Qy 2188 TGCTCGAAAGAAATACAGAACTAGAAAGTGGCTAAGCGAGGAAATGGAGGTACAAAATCA 2247  
Db 2161 TGCTCGAAAGAAATACAGAACTAGAAAGTGGCTAAGCGAGGAAATGGAGGTACAAAATCA 2220  
Qy 2248 ACATGCAAAAAGAGAGTCTCTTGTGATGATCTTTTATAGATCAATCTCTTTTAAAAAG 2307  
Db 2221 ACATGCAAAAAGAGAGTCTCTTGTGATGATCTTTTATAGATCAATCTCTTTTAAAAAG 2280  
Qy 2308 GAGAGATAA CTGAGGATTTTAAAAAGGCCATCGAAAAAATCTCTTAGTAGCATCTAC 2367  
Db 2281 GAGAGATAA CTGAGGATTTTAAAAAGGCCATCGAAAAAATCTCTTAGTAGCATCTAC 2340  
Qy 2368 TTCAGGCCAACAGGTTATATGAATATATAGTATAGAGCGATTTAAGTTTCAATGTT 2427  
Db 2341 TTCAGGCCAACAGGTTATATGAATATATAGTATAGAGCGATTTAAGTTTCAATGTT 2400  
Qy 2428 TTAATGGCCCTAAATTTTAAATAAATGCAAAAACTTTGATCTTTTGTATGTAACAAAT 2487  
Db 2401 TTAATGGCCCTAAATTTTAAATAAATGCAAAAACTTTGATCTTTTGTATGTAACAAAT 2460  
Qy 2488 TGTGTTGCTGTTTTTTCAGGCTTTGTCATGTCATCTTTTTCATTTTAAATGCTGTTTG 2547  
Db 2461 TGTGTTGCTGTTTTTTCAGGCTTTGTCATGTCATCTTTTTCATTTTAAATGCTGTTTG 2520







```
; NUMBER OF SEQ ID NOS: 371
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 343
; LENGTH: 646
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-401-064-343

Query Match      12.5%; Score 552.2; DB 4; Length 646;
Best Local Similarity 93.0%; Pred. No. 4.5e-131;
Matches 609; Conservative 4; Mismatches 32; Indels 10; Gaps 3;

QY 3419 AAAAGCTCTCTAGCCTTCCTAGGACAGAAATGGGTGAGCAATCTGCAATTTCTACT 3478
Db 645 AAAGCTTTCTAGCCTTCCTAGGACAGAAATGGGTGAGCAATCTGCAATTTCTACT 586

QY 3479 ACAGGCATTGAGACCAAGTTAGATTATTGAAATATTATAGAGAGTTATGAACACTTAAAT 3538
Db 585 ACAGGCATTGAGACCAAGTTAGATTATTGAAATATTATAGAGAGTTATGAACACTTAAAT 526

QY 3539 ATGATAGTGGTATGACATTTGGATAGAACATGGGATCTTTAGAAAGTAAATTCACAGGCG 3598
Db 525 ATGATAGTGGTATGACATTTGGATAGAACATGGGATCTTTAGAAAGTAAATTCACAGGCG 466

QY 3599 ATATTAGTTGATGAAATGAGTCATTTGAGTCTTTTAAATAGCCATGTATCATAAATTACCA 3658
Db 465 ATATTAGTTGATGAAATGAGTCATTTGAGTCTTTTAAATAGCCATGTATCATAAATTACCA 406

QY 3659 AGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTTAAGGAATATATGACAGATT 3718
Db 405 AGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTTAAGGAATATATGACAGATT 346

QY 3719 AATATTGTTCTGTCATGCCACAAATCCCTTTTAAAGAAAGACTGCCCTACTATAGCAG 3778
Db 345 AATATTGTTCTGTCATGCCACAAATCCCTTTTAAAGAAAGACTGCCCTACTATAGCAG 286

QY 3779 TTTTATATTGTCATTTATGAATTAATGAATGAGGAGTTCTGGTACTCTCTGCTTT 3838
Db 285 TTTTATATTGTCATTTATGAATTAATGAATGAGTGA- GAGTCTGGTACTCTCTGCTTT 227

QY 3839 ACAAAATATTGGGTGTTGTCAGTATTTTCCCTTTTAAACMTTCCCAATTCGGGTGCT 3898
Db 226 ACAAAATATTGG- - -TGTGTGAGTATTTTCTTTTAAACCAATTCCAATCGGTGTAG 171

QY 3899 AGGTGATGTTCCATTTGGTTTAAATTTGATATATCCCTGATAGCTATAATTTGGTCA 3958
Db 170 TGATG- - -TTTCATTTTGGTTTAAATTTGATATATCCCTGATAGCTATAATTTGGTCA 116

QY 3959 AGAAATTTCTTTATACATTTAGATGCAAGTCTCTTGGCGGATATAGTATTGAGATATTA 4018
Db 115 AGAAATTTCTTTATACATTTAGATGCAAAATCTCTTGGCGGATATATGATTGAGATATTA 56

QY 4019 CACCTAGTCTGGGCTTGACTGTTTCTTTTATGCTCTTTTATGATGAATPAGAAGTTTT 4073
Db 55 CACCTAGTCTGGGCTTGACTGTTTCTTTTATGCTCTTTTATGATGAATPAGAAGTTTT 1

RESULT 3
US-09-949-016-79057
; Sequence 79057, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14033
; LENGTH: 784019
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(784019)
; OTHER INFORMATION: n = A,T,C or G

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 79057
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-79057

Query Match      1.9%; Score 84; DB 4; Length 601;
Best Local Similarity 56.6%; Pred. No. 2.9e-11;
Matches 248; Conservative 3; Mismatches 153; Indels 34; Gaps 4;

QY 3945 TATATTTGGTCATAGAAATCTTTATACATCTTAGAGTCAAGTCTCTTGYGGGATATAC 4004
Db 60 TAATATTTGAAATTAAGGAATCTATATATATTTTGGATACATCTTTTTTTTTCATATGT 119

QY 4005 GTATTGAGATATTACACCTAGTCTGTGGCTTGACTGTTTCTTTATGCTCTTTTATGATGA 4064
Db 120 ACTGTGA-ATTTCTCTCTTGTCTATTTATTCCTTAATGTAATCTTTTGAAGAGC 178

QY 4065 AGAAGTTTAAATTTTGAAGGTCAAATTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTCT 4124
Db 179 AGAAGTTTAAATTAAGTCTAGTGTCCAAATTTATCTTTTCTTTTCTTTTCTTTTCTTTCT 238

QY 4125 CTCC- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - 4154
Db 239 TTCTAAGAACCTTTACCAATCACCAGGTCACAAAATATTTTCTTATGTTTCTTTCTAGA 298

QY 4155 CTGCTCTATTATATAAACTTTTATATTTTATATTTTGTG- - - - - - - - - - - - 4212
Db 299 CASTCCATAAATCTTCAATCTTTACATTCAGATCCTATGTTCCATCTCAAAATTAATTTTGA 358

QY 4213 TATGTTGTAATTAAGGATCAGGTTCTTTTTCCTCCCAACAAGTATCCAGTCAATGT 4272
Db 359 TATGTTGTAAGTAAATGAGTGTAGAGGTTTATTTTCTCTCCATAAGACTCTCCAAATTTCT 418

QY 4273 AACACTGTTTATGAAGAATTTATCTTTTCTTCTTAAATTAACCTTTGCCA-ATTAGTAA 4331
Db 419 AGCATATTTGTTCAAGACTAACTCTCGGCTATTCATTTCTTTGGCATCTTTTGTCAA 478

QY 4332 AAATCAATTAACCATRMA 4349
Db 479 AAATCAGTTGACTATATA 496

RESULT 4
US-09-949-016-14033
; Sequence 14033, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14033
; LENGTH: 784019
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(784019)
; OTHER INFORMATION: n = A,T,C or G
```



US-09-949-016-14033

Query Match 1.9%; Score 82.8; DB 4; Length 784019;  
Best Local Similarity 56.6%; Pred. No. 3.6e-09;  
Matches 248; Conservative 2; Mismatches 154; Indels 34; Gaps 4;  
  
QY 3945 TATAATGGGTCATAGAAATCTTTATACATTTCTAGATGAAGTCTCTTGYCGGATATAC 4004  
DB 467327 TAATATTGAATTATAGAAATCTTATATATTTTGGATACATTTCTTTTTCATATGT 467386  
  
QY 4005 GTATTGAGATATTACACCTAGTCTGGCTTGACTGTTTCTTTATGTCTTTTTCATATGT 463519  
DB 463520 ACTGTGA-ATTTTCTCTTTGCTTGTCTATTATTTCTTAAATGGTATCTTTTGAAGC 463578  
  
QY 4065 AGAAGTTTTAAATTTTGAACAAGTCAAAATTTATTTTCTTTTGTGATATTTTCTTCT 4124  
DB 463579 AGAAGTTATAAATACTGATAGTGTCCAAATTTATCTTTTCAATTTGTTAGTCTTTCTTG 463638  
  
QY 4125 CTCC-----AATTTAACCCCAAGATTTCCAGATATT 4154  
DB 467387 ACTGTGA-ATTTTCTCTTTGCTTGTCTATTATTTCTTAAATGGTATCTTTTGAAGC 467445  
  
QY 4065 AGAAGTTTTAAATTTTGAACAAGTCAAAATTTATTTTCTTTTGTGATATTTTCTTCT 4124  
DB 467446 AGAAGTTATAAATACTGATAGTGTCCAAATTTATCTTTTCAATTTGTTAGTCTTTCTTG 467505  
  
QY 4125 CTCC-----AATTTAACCCCAAGATTTCCAGATATT 4154  
DB 467506 TTCTAAGAACCTTTTACCAATCACCAGGTCAAAAAATATTTCTCTATGTTTCTTCTAGA 467565  
  
QY 4155 CTGCTCTATTATATAAATCTTTATATTTTATTTTGTG--ATCTACCTTGAATTTGATG 4212  
DB 467566 CACTCATAACTTCAATCTTTTACATTCAGATCTATGTTCCATCTCAAAATTAATTTTGTGA 467625  
  
QY 4213 TATGTTGTGAATTTATGGATCAGGTTCTTTTTCCTCCCATACAAGTATCCAGTCAATGT 4272  
DB 467626 TATGGTGTGAGGTAATGGTAGAGTTTATTTTCTCTCCATAGACTCTCCAATTTCTTCT 467685  
  
QY 4273 AACACTGTTTATTGAAGAATTAATCTTCTTCTCATTAAATTTACCTTGCCA-ATTAGTAA 4331  
DB 467686 AGCACTATTGTTTCAAGACTAATCTCGGCTATTCAATTTCTTGGCATCTTTGTCAA 467745  
  
QY 4332 AATCAATTAACCATRMA 4349  
DB 467746 AATCAGTTGACTATATA 467763

RESULT 5

US-09-949-016-12777  
; Sequence 12777, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12777  
; LENGTH: 828152  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(828152)  
; OTHER INFORMATION: n = A, T, C or G

US-09-949-016-12777

Query Match 1.9%; Score 82.8; DB 4; Length 828152;  
Best Local Similarity 56.6%; Pred. No. 3.7e-09;  
Matches 248; Conservative 2; Mismatches 154; Indels 34; Gaps 4;  
  
QY 3945 TATAATGGGTCATAGAAATCTTTATACATTTCTAGATGAAGTCTCTTGYCGGATATAC 4004

DB 463460 TAATATTGAATTATAGAAATCTTATATATATTTTGGATACATTTCTTTTTCATATGT 463519  
QY 4005 GTATTGAGATATTACACCTAGTCTGGCTTGACTGTTTCTTTATGTCTTTTGTGATGAAT 4064  
DB 463520 ACTGTGA-ATTTTCTCTTTGCTTGTCTATTATTTCTTAAATGGTATCTTTTGAAGC 463578  
  
QY 4065 AGAAGTTTTAAATTTTGAACAAGTCAAAATTTATTTTCTTTTGTGATATTTTCTTCT 4124  
DB 463579 AGAAGTTATAAATACTGATAGTGTCCAAATTTATCTTTTCAATTTGTTAGTCTTTCTTG 463638  
  
QY 4125 CTCC-----AATTTAACCCCAAGATTTCCAGATATT 4154  
DB 463639 TTCTAAGAACCTTTTACCAATCACCAGGTCAAAAAATATTTCTCTATGTTTCTTCTAGA 463698  
  
QY 4155 CTGCTCTATTATATAAATCTTTATATTTTATTTTGTG--ATCTACCTTGAATTTGATG 4212  
DB 463699 CACTCATAACTTCAATCTTTTACATTCAGATCTATGTTCCATCTCAAAATTAATTTTGTGA 463758  
  
QY 4213 TATGTTGTGAATTTATGGATCAGGTTCTTTTTCCTCCCATACAAGTATCCAGTCAATGT 4272  
DB 463759 TATGGTGTGAGGTAATGGTAGAGTTTATTTTCTCTCCATAGACTCTCCAATTTCTTCT 463818  
  
QY 4273 AACACTGTTTATTGAAGAATTAATCTTCTTCTCATTAAATTTACCTTGCCA-ATTAGTAA 4331  
DB 463819 AGCACTATTGTTTCAAGACTAATCTCGGCTATTCAATTTCTTGGCATCTTTGTCAA 463878  
  
QY 4332 AATCAATTAACCATRMA 4349  
DB 463879 AATCAGTTGACTATATA 463896

RESULT 6

US-09-573-080A-425/c  
; Sequence 425, Application US/09573080A  
; Patent No. 6828097  
; GENERAL INFORMATION:  
; APPLICANT: JOAN, KNOLL  
; APPLICANT: ROGAN, PETER  
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERAT  
; FILE REFERENCE: 30307  
; CURRENT APPLICATION NUMBER: US/09/573,080A  
; CURRENT FILING DATE: 2000-05-16  
; NUMBER OF SEQ ID NOS: 479  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 425  
; LENGTH: 3285  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: repeat region  
; LOCATION: (1)...(3285)  
; OTHER INFORMATION: lme\_orf2  
; NAME/KEY: misc feature  
; OTHER INFORMATION: n is a, c, g or t  
; PUBLICATION INFORMATION:  
; PUBLICATION INFORMATION:  
; AUTHORS: Jurka, J; Walichewicz, J; Milosavljevic, A  
; TITLE: Prototypic sequences for human repetitive DNA  
; JOURNAL: Journal of Molecular Evolution  
; VOLUME: 35  
; ISSUE: 4  
; PAGES: 286-291  
; DATE: 1992-10-  
; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)  
; DATABASE ENTRY DATE: 1996-01-26  
; DATABASE ENTRY DATE: 1996-01-26  
; US-09-573-080A-425

Query Match 1.9%; Score 81.6; DB 4; Length 3285;  
Best Local Similarity 58.8%; Pred. No. 3.2e-10;  
Matches 284; Conservative 7; Mismatches 137; Indels 55; Gaps 7;







US-09-949-016-14061

Query Match 1.8%; Score 78; DB 4; Length 75216;  
Best Local Similarity 56.4%; Pred. No. 1.1e-08;  
Matches 229; Conservative 1; Mismatches 161; Indels 15; Gaps 4;  
  
QY 3945 TATAATGGGTGATAGAAATCTTTATACATCTAGATGCAAGTCTCTTGYCGGATATAC 4004  
DB 796 TATATTGAGTTGTAAGAGTCTTTTACATCTAGATAAAGGTCCCTTATCAGATATGC 737  
  
QY 4005 GTATTGAGATATTACACCTAGTCTGTGGCTTGACTGTTTCTTATGCTCTTTTGTGATGAAT 4064  
DB 736 TT-ATCAATATATATCTATAGAGTCTTTTAC----TTTCTGGTGTCTTTTAAACAC 681  
  
QY 4065 AGAAGTTTAAATTTTGACAAGGTCAAAATTAATTT-----TTTCTTTTGTGTTGAT 4115  
DB 680 AGAAGTTTAAATTTTGATGAAGTCCAAATTAATCTGTTTGTGTTGTTGTTGTTGTTGTT 621  
  
QY 4116 ATTTTCTCTCCAAATTTAAACCCAGATTTTCAAGATTTCTGCTCTATATATATAAACTTT 4175  
DB 620 GCTTTTGGTGCATATCCCAAGAATCCCTTAACCTTGTGTTCTTTTCCAAAGAGTTTAACTG 561  
  
QY 4176 ATATTTTATATTGATGATCTACTTGAATTTGATATGATGTTGTGAAT--TATGGATCA 4233  
DB 560 ATCAATTTAGTCTGTGATCCATTTTGAGTTAAATTTTCTTGTAGCATGGTATGAGAGT 501  
  
QY 4234 GGGTCTCTTTTCCCCCATACAAAGTATCCAGTCAITTTGAACACTGTTTATTGAAAGAAT 4293  
DB 500 CCAACCTCATTTTTCGATGTCTATCCAGTGTCCCAACATCTTGTGGAAGAC 441  
  
QY 4294 TATCTTTCTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4339  
DB 440 TCCTCTTTCTCCACAGAATGGTCTGTGGACCTGTCTAGACTGAAT 395

RESULT 9

US-09-949-016-12055  
; Sequence 12055, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12055  
; LENGTH: 141454  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(141454)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12055

Query Match 1.7%; Score 75.4; DB 4; Length 141454;  
Best Local Similarity 55.1%; Pred. No. 1.1e-07;  
Matches 212; Conservative 1; Mismatches 162; Indels 10; Gaps 3;  
  
QY 3945 TATAATGGGTGATAGAAATCTTTATACATCTAGATGCAAGTCTCTTGYCGGATATAC 4004  
DB 126363 TATATTGAATCTTACTGTTCTTCACTGCTCTGGATACAGTCTCTTGTGACACAT 126422  
  
QY 4005 GTATTG--AGATATTACACCTAGTCTGTGGCTTGAC-----TGTTTTCTTTATGCTTTT 4056

DB 126423 GTTTTGCATAATCTTTTCCCAAGTCTGTTGCATGACAATTTGTTTTTCTGTAAGTGT 126482  
QY 4057 TGATGAATAGAAGTTTAAATTTTGACAAGTCAAAATTTATTTTCTTTCTTTGATA 4116  
DB 126483 TAATGACGAGAGTCTCTCATTTTATGAGATGTAATTTATTAACCTCTCTTTAATGAT 126542  
QY 4117 TTTTCTCTCCAAATTTAAACCCAGATTTCAAGATTTCTGCTCTATATATAAACTTTA 4176  
DB 126543 TATTAATTTCTGTGTTCTAAGAAAGCTTTGCTTACCATAGTCAGGAAGTTTCCCTTTG 126602  
QY 4177 TATTTTATATTGTCATCTACCTTCAATTTGAT--ATGATGTTGTAATTTGGAATCAG 4234  
DB 126603 TTTAGCCCTATAGTCCATCTCTTTTCTTTTTCATGAATGGTGTAGTAGAAATTTGA 126662  
QY 4235 GGTCTCTTTTTCCTCCCATACAAAGTATCCAGTCAITTTGAACACTGTTTATTGAAAGAAT 4294  
DB 126663 AGTTCAGTTTCTCTTATACCCAGATTCAGTGTTCAGCACAAATTAATTTGGAAGAAT 126722  
QY 4295 ATCTTTCTCTCAATTAATTAATTAATTTACCTTG 4319  
DB 126723 TTTACTTTCTCATTAATGTTTCTCTG 126747

RESULT 10

US-09-949-016-17190  
; Sequence 17190, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17190  
; LENGTH: 141455  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(141455)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-17190

Query Match 1.7%; Score 75.4; DB 4; Length 141455;  
Best Local Similarity 55.1%; Pred. No. 1.1e-07;  
Matches 212; Conservative 1; Mismatches 162; Indels 10; Gaps 3;  
  
QY 3945 TATAATGGGTGATAGAAATCTTTATACATCTAGATGCAAGTCTCTTGYCGGATATAC 4004  
DB 126363 TATATTGAATCTTACTGTTCTTCACTGCTCTGGATACAGTCTCTTGTGACACAT 126422  
QY 4005 GTATTG--AGATATTACACCTAGTCTGTGGCTTGAC-----TGTTTTCTTTATGCTTTT 4056  
DB 126423 GTTTTGCATAATCTTTTCCCAAGTCTGTTGCATGACAATTTGTTTTTCTGTAAGTGT 126482  
QY 4057 TGATGAATAGAAGTTTAAATTTTGACAAGTCAAAATTTATTTTCTTTCTTTGATA 4116  
DB 126483 TAATGACGAGAGTCTCTCATTTTATGAGATGTAATTTATTAACCTCTCTTTAATGAT 126542  
QY 4117 TTTTCTCTCCAAATTTAAACCCAGATTTCAAGATTTCTGCTCTATATATAAACTTTA 4176  
DB 126543 TATTAATTTCTGTGTTCTAAGAAAGCTTTGCTTACCATAGTCAGGAAGTTTCCCTTTG 126602



```
QY 4177 TATTTTATATTTGATCTACCTGAAATTGAT--ATGTATGTTGTGAATTTGATGATCAG 4234
Db 126603 TTTAGCCCTATAGTCACATCTTTCTTTTTCATGAATGGTGTGTAGTAAATTTCA 126662
QY 4235 GGTTCCTTTTTCCTCCCATACAGATGATCCAGTCATTTGAACACTGTTTATTTGAAAGAAAT 4294
Db 126663 AGTTCAGTTTCTCTTTCATACCGATATCCAGTTGTTCCAGCACAAATTAATTCGAAAGATT 126722
QY 4295 ATCTTTCCTCATTAATAATACCTTG 4319
Db 126723 TTTACTTTCTCATTAATGTTTCTCTG 126747
```

```
RESULT 11
US-09-949-016-13379/c
; Sequence 13379, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13379
; LENGTH: 173992
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13379
```

```
Query Match 1.7%; Score 74.8; DB 4; Length 173992;
Best Local Similarity 57.3%; Pred. No. 1.7e-07;
Matches 172; Conservative 1; Mismatches 123; Indels 4; Gaps 2;

QY 3839 ACAAAATATTTGGGTGTGTCAGATATTTTCCCTTTTAAACCMTTCCTCCAAATTCGGGTGTGT 3898
Db 144215 ACTAAGGTGTTGATAGTCTTTTCATGTGTAATTTGCCACCTACACAATCTTCTCGATG 144156
QY 3899 AGGTGATGTTCCATTTGGGTTTTTAATTTGTATATCCCTGATAGCTATAAATTTGGGTGAT 3958
Db 144155 AACTGTCTGTTCAAATTTTGGCCATTTTAAATGCAGTTGTTGTTTCTTGAGTTT 144096
QY 3959 AGNAATCTTTATACATTTAGATGCAAGTCCTTGYCGGATATACGTATTGAGATATTA 4018
Db 144095 AAGAATCTTTATACATTTCTGGATACAGTCCTTCAGACACATATGATTTAAATTTTA 144036
QY 4019 CACTAGTCTGTGGCTTGACTGT---TTTCTTTATGCTTTTGTATGAATGAAGTTTAA 4075
Db 144035 -TCCAGTCTGTGGCTTGCTTTTCAATCTCTTAAAGTTTGAAGACGAGAAATCTTA 143977
QY 4076 ATTTGACAGGTCAAATTTATTTTCTTTTGTGTTGATATTTTCTCTCAATTTAA 4135
Db 143976 ATTTGATGAAGTCCAATTTATTAATGTTTATGGATTGCACCTTTTGTGCATATCTAA 143917
```

```
RESULT 12
US-09-902-540-1357/c
; Sequence 1357, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
```

```
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1357
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(612)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1357

Query Match 1.6%; Score 69.6; DB 4; Length 612;
Best Local Similarity 47.9%; Pred. No. 1.4e-07;
Matches 224; Conservative 1; Mismatches 242; Indels 1; Gaps 1;

QY 3779 TTTTATATTTGTCAAATTTATGAATATGAATGAGGAGTTCTGGTACCTCCTGCTTT 3838
Db 504 TTTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 445
QY 3839 ACAAAATATTTGGGTGTGTCAGATATTTTCCCTTTTAAACCMTTCCTCCAAATTCGGGTGTGT 3898
Db 444 TTTAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 385
QY 3899 AGGTGATGTTTCCATTTGGGTTTTAAATTTGTATATCCCTGATAGCTATAAATTTGGGTGAT 3958
Db 384 ATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 326
QY 3959 AGAAATCTTTATACATTTAGATGCAAGTCCTTGYCGGATATACGTATTGAGATATTA 4018
Db 325 TTTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 266
QY 4019 CACTAGTCTGTGGCTTGACTGT---TTTCTTTATGCTTTTGTATGAATGAAGTTTAAATTT 4078
Db 265 ATTTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 206
QY 4079 TTGACAAAGTCAAATTTATTTTCTTTTGTGATATTTTCTCTCTCAATTTAACC 4138
Db 205 TTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 146
QY 4139 CAAGATTTTCAGATATTTCTGCTCTATTATATAAACTTTATTTTCTTTTCTTTTCTTTTCTTTT 4198
Db 145 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 86
QY 4199 CTTGAATTGATATGATGTGTGTAATTTATGGAATTCAGGTTCTTTTCTTTTCTTTTCTTTTCTTTT 4246
Db 85 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 38
```

```
RESULT 13
US-09-949-016-15784/c
; Sequence 15784, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15784
```







Db 126 A 126

Search completed: August 30, 2005, 15:39:18  
Job time : 684 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2005, 05:17:43 ; Search time 2560 Seconds  
(without alignments)  
11253.603 Million cell updates/sec

Title: US-09-837-138-1  
Perfect score: 4403  
Sequence: 1 ttcgcagcagcgcggttg.....accgcggtgagctccagct 4403

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7331713 seqs, 3271544945 residues

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_NA:\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*  
19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq:\*  
20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq:\*  
21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq:\*  
22: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq:\*  
24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*  
25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	564.6	12.8	622	19 US-10-474-495-34	Sequence 34, Appl
2	552.2	12.5	646	9 US-09-922-217-343	Sequence 343, App
3	552.2	12.5	646	9 US-09-813-263-343	Sequence 343, App
4	552.2	12.5	646	13 US-10-025-380-343	Sequence 343, App
5	544.8	12.4	561	11 US-09-969-034-1463	Sequence 1463, Ap
6	537.4	12.2	541	9 US-09-878-178-1737	Sequence 1737, Ap
7	537.4	12.2	541	13 US-10-046-935-1737	Sequence 1737, Ap

8	537.4	12.2	541	14 US-10-146-502-1737	Sequence 1737, Ap
9	507	11.5	587	15 US-10-102-524-532	Sequence 532, App
10	478.4	10.9	481	17 US-10-242-535A-26197	Sequence 26197, A
11	478.4	10.9	481	17 US-10-085-783A-26197	Sequence 26197, A
12	438.2	10.0	439	17 US-10-242-535A-7179	Sequence 7179, Ap
13	438.2	10.0	439	18 US-10-085-783A-7179	Sequence 7179, Ap
14	405.6	9.2	468	10 US-09-918-995-13834	Sequence 13834, A
c	362	8.2	483	11 US-09-969-034-1974	Sequence 1974, A
15	258.4	5.5	360	9 US-09-783-590-3449	Sequence 3449, Ap
16	240.8	5.5	267	20 US-10-425-115-120439	Sequence 120439,
17	239.4	5.4	272	9 US-09-923-876-5167	Sequence 5167, Ap
18	239.4	5.4	272	10 US-09-923-876-5167	Sequence 5167, Ap
c	212.6	4.8	475	9 US-09-917-800A-803	Sequence 803, App
21	191	4.3	404	9 US-09-783-590-8929	Sequence 8929, Ap
22	167.6	3.8	315	9 US-09-960-352-8727	Sequence 8727, Ap
23	121.8	2.8	540	21 US-10-867-578-1	Sequence 1, Appli
24	119.4	2.7	442	9 US-09-864-761-11287	Sequence 11287, A
25	114	2.6	114	9 US-09-864-761-27987	Sequence 27987, A
26	97.6	2.2	6237	15 US-10-311-455-331	Sequence 331, App
c	92.4	2.1	6237	15 US-10-311-455-332	Sequence 332, App
28	81.6	1.9	3285	10 US-09-854-867-425	Sequence 425, App
c	81.6	1.9	3285	21 US-10-786-970A-425	Sequence 425, App
30	80	1.8	1300	19 US-10-668-749A-1	Sequence 1, Appli
31	74.6	1.7	1980090	20 US-10-719-993-6815	Sequence 6815, Ap
32	74.6	1.7	1980090	21 US-10-741-600-17676	Sequence 17676, A
c	73.6	1.7	539	14 US-10-198-846-1369	Sequence 1369, Ap
34	72.2	1.6	7442	18 US-10-221-714A-409	Sequence 409, App
35	72	1.6	9539	14 US-10-239-676-52	Sequence 52, Appl
36	72	1.6	9539	15 US-10-240-453-54	Sequence 54, Appl
c	71.4	1.6	739	20 US-10-425-115-55020	Sequence 55020, A
38	71.4	1.6	1062	21 US-10-425-115-120013	Sequence 120013,
c	71.2	1.6	322885	20 US-10-679-366-3	Sequence 3, Appli
40	71	1.6	14006	15 US-10-311-455-1931	Sequence 1931, Ap
c	70	1.6	6056	15 US-10-311-455-753	Sequence 753, App
c	70	1.6	6056	18 US-10-221-714A-85	Sequence 85, Appl
43	69.8	1.6	1243	20 US-10-425-115-172717	Sequence 172717,
44	69.6	1.6	15548	15 US-10-311-455-2128	Sequence 2128, Ap
45	69.6	1.6	3673778	16 US-10-312-841-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-10-474-495-34  
; Sequence 34, Application US/10474495  
; Publication No. US20040171003A1  
; GENERAL INFORMATION:  
; APPLICANT: YOSHIKAWA, Yoshie et al.  
; TITLE OF INVENTION: CANCER-ASSOCIATED GENES  
; FILE REFERENCE: 1422-0606P  
; CURRENT APPLICATION NUMBER: US/10474,495  
; PRIOR FILING DATE: 2003-10-09  
; PRIOR APPLICATION NUMBER: JP 2001-112039  
; PRIOR FILING DATE: 2001-04-10  
; PRIOR APPLICATION NUMBER: JP 2001-290193  
; PRIOR FILING DATE: 2001-09-21  
; NUMBER OF SEQ IDS: 264  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 34  
; LENGTH: 622  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: variation  
; LOCATION: 553, 597  
; OTHER INFORMATION: n is a or c or g or t.  
US-10-474-495-34

Query Match 12.8%; Score 564.6; DB 19; Length 622;  
Best Local Similarity 97.6%; Pred. No. 1.1e-113;  
Matches 604; Conservative 0; Mismatches 11; Indels 4; Gaps 3;



QY 840 TTTTGGCTCGGGAACGTGTGTTGATACAGGAATACAACTCACAGACTTAATTC 899  
Db 1 TTTTGGCTCGGGAACGTGTGTTGATACAGGAATACAACTCACAGACTTAATTC 60  
QY 900 CTGACTGTGTCAGAGAAATGGATTGATCAATATGATATGCTCCAAAGGCAAGGTCTTA 959  
Db 61 CTGACTGTGTCAGAGAAATGGATTGATCAATATGATATGCTCCAAAGGCAAGGTCTTA 120  
QY 960 GACCTATTCCTGAAGCAGAAATGGATTGCGGTGATTTTCATGACTACAAAGAAATTA 1019  
Db 121 GACCTATTCCTGAAGCAGAAATGGATTGCGGTGATTTTCATGACTACAAAGAAATTA 180  
QY 1020 GTGATCCTCAGGCGCAATCCAGTACAGGAATTAAGACACACACTCCAGGACCAAGCCTTT 1079  
Db 181 GTGATCCTCAGGCGCAATCCAGTACAGGAATTAAGACACACACTCCAGGACCAAGCCTTT 240  
QY 1080 CACAAGCGGTGTCAGTTGATGAAAACTAATGCCAAGCGCCCGAGTGAACACTACAAAT 1139  
Db 241 CACAAGCGGTGTCAGTTGATGAAAACTAATGCCAAGCGCCCGAGTGAACACTACAAAT 300  
QY 1140 AGTAGCTGACACAGAAATCAGAGCAAGCAGATACATGGGATTTGAGTGAAGGCCAAAAG 1199  
Db 301 AGTAGCTGACACAGAAATCAGAGCAAGCAGATACATGGGATTTGAGTGAAGGCCAAAAG 360  
QY 1200 AATCAAGCTCTCCAAATGGAACAAATTCAGATGCTTTCACAAAGCGCACCACTG 1259  
Db 361 AATCAAGCTCTCCAAATGGAACAAATTCAGATGCTTTCACAAAGCGCACCACTG 420  
QY 1260 TAAAGGAGTCTCTGCAAAACAAAGCTCTAATAATAGTATGATGCTCAAAATACCTTTGGCTA 1319  
Db 421 TAAAGGAGTCTCTGCAAAACAAAGCTCTAATAATAGTATGATGCTCAAAATACCTTTGGCTA 480  
QY 1320 AGATGAGATCCCAAACTACAGCTTTCCAACTAAATGCGAAGTGAATATAAAGTA 1379  
Db 481 AGATGAGATCCCAAACTACAGCTTTCCAACTAAATGCGAAGTGAATATAAAGTA 540  
QY 1380 AGATGAGGCTTC-CACAGCAGCAGCAGACCAAC-TCCATCAGAACTACTTT--CAGCG 1435  
Db 541 AGATGAGGCTTTTTCAGCAGCAGCAGCAGCAACTTCCATCAGAACTACTTTTCAGNCCG 600  
QY 1436 TCTTACCAAAAAAGGAAA 1454  
Db 601 TCTACCGGAAAAA 619

## RESULT 2

US-09-922-217-343/c  
; Sequence 343, Application US/09922217  
; Patent No. US20020076414A1

## GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Meagher, Madeleine Joy  
; APPLICANT: Stolk, John A.  
; APPLICANT: Wang, Tongfong  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Smith, Carole Lynn  
; APPLICANT: King, Gordon E.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Clapper, Jonathan D.

## TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS

; FILE REFERENCE: 210121.471C13

; CURRENT APPLICATION NUMBER: US/09/922,217

; NUMBER OF SEQ ID NOS: 1124

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 343

; LENGTH: 646

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-922-217-343

## Query Match

Best Local Similarity 12.5%; Score 552.2; DB 9; Length 646;

Matches 609; Conservative 4; Mismatches 32; Indels 10; Gaps 3;

QY 3419 AAAAGCTTCTCAGCCTTCCTAGGGAACAGAAATGGGTGAGCCAACTCGCAATTTCTACT 3478  
Db 645 AAAAGCTTCTCAGCCTTCCTAGGGAACAGAAATGGGTGAGCCAACTCGCAATTTCTACT 586  
QY 3479 ACAGGCATTGAGACCAAGTTAGATTATGAAATATTTATAGAGAGTTATGAAACACTTAAAT 3538  
Db 585 ACAGGCATTGAGACCAAGTTAGATTATGAAATATTTATAGAGAGTTATGAAACACTTAAAT 526  
QY 3539 ATGATAGTGGTATGACATTTGGATAGAAACATGGGATACCTTTAGAGTGAATTTGACAGGCG 3598  
Db 525 ATGATAGTGGTATGACATTTGGATAGAAACATGGGATACCTTTAGAGTGAATTTGACAGGCG 466  
QY 3599 ATATTAGTTGATGAAATGGAGTCAATTTGAGTCTCTTTATAGCCATGATATCAATATACCA 3658  
Db 465 ATATTAGTTGATGAAATGGAGTCAATTTGAGTCTCTTTATAGCCATGATATCAATATACCA 406  
QY 3659 AGTGAAGCTGGTGAACATATGCTCTCCATTTTACAGTTAAGGAATAATAATGGAAGATTT 3718  
Db 405 AGTGAAGCTGGTGAACATATGCTCTCCATTTTACAGTTAAGGAATAATAATGGAAGATTT 346  
QY 3719 AATATTGTTTCTGTCATGCCCAATCCCTTTCTAAGGAAGACTGCCCTACTATAGCAG 3778  
Db 345 AATATTGTTTCTGTCATGCCCAATCCCTTTCTAAGGAAGACTGCCCTACTATAGCAG 286  
QY 3779 TTTTATATTTGTCAAATTTATGAATATATGAATGAGGAGTTCTGGTACCTCTCTCTTT 3838  
Db 285 TTTTATATTTGTCAAATTTATGAATATATGAATGAGTCAAGTCTGGTACCTCTCTCTTT 227  
QY 3839 ACAATAATGGGTGTTGTCAGTATTTTCCCTTTTAAACMTTCCCAATTCGGGTGTGT 3898  
Db 226 ACAATAATGG---TGTGTCAGTATTTTCCCTTTTAAACMTTCCCAATTCGGGTGTGT 171  
QY 3899 AGTGGATGTTTCCATTTGGGTTTTAAATTTGATATATCCCTGATAGCTATTAATTTGGTCA 3958  
Db 170 TGATG-----TTTCATTTTGGTTTTAAATTTGTATATCCCTGATAGCTATAATTTGGTCA 116  
QY 3959 AGAAATCTTTATATACATTTAGATGCAAGTCTCTTGYCGGATATATACGTATTTGAGATATTA 4018  
Db 115 AGAAATCTTTATATACATTTAGATGCAAGTCTCTTGYCGGATATATATGATTTGAGATATTA 56  
QY 4019 CACCTAGTCTGTGGCTTGACTGTTTCTTTATGCTTTTGTGATGAATAGAAGTTTT 4073  
Db 55 CACCTAGTCTGTGGCTTGACTGTTTCTTTATGCTTTTGTGATGAATAGAAGTTTT 1

## RESULT 3

US-09-833-263-343/c

; Sequence 343, Application US/09833263

; Patent No. US20020110547A1

## GENERAL INFORMATION:

; APPLICANT: Wang, Aijun  
; APPLICANT: Clapper, Jonathan D.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Meagher, Madeleine J.  
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.471C12  
; CURRENT APPLICATION NUMBER: US/09/833,263  
; CURRENT FILING DATE: 2001-04-10  
; NUMBER OF SEQ ID NOS: 1093  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 343  
; LENGTH: 646  
; TYPE: DNA  
; ORGANISM: Homo sapien

US-09-833-263-343



```
Query Match      12.5%; Score 552.2; DB 9; Length 646;
Best Local Similarity 93.0%; Pred. No. 5.8e-111;
Matches 609; Conservative 4; Mismatches 32; Indels 10; Gaps 3;

QY 3419 AAAAGCTTCTCAGCCTTCTAGGGAACAGAAATTTGGGTGAGCCCAATCTGCAATTTCTACT 3478
Db 645 AAAGCTTCTCAGCCTTCTAGGGAACAGAAATTTGGGTGAGCCCAATCTGCAATTTCTACT 586

QY 3479 ACAGGCATTGAGACCAAGTTAGATTATTGAATATTATAGAGAGTTATGAACACTTAATTT 3538
Db 585 ACAGGCATTGAGACCAAGTTAGATTATTGAATATTATAGAGAGTTATGAACACTTAATTT 526

QY 3539 ATGATAGTGTATGACATTCGATAGAACATGGGATCTTTAGAAAGTAGAATTCACAGGGC 3598
Db 525 ATGATAGTGTATGACATTCGATAGAACATGGGATCTTTAGAAAGTAGAATTCACAGGGC 466

QY 3599 ATATTAGTTGATGAAATGGAGTCATTTGAGTCTCTTTTAATAGCCATGTATCATTAATTTACCA 3658
Db 465 ATATTAGTTGATGAAATGGAGTCATTTGAGTCTCTTTTAATAGCCATGTATCATTAATTTACCA 406

QY 3659 AGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGGAATATATAGGACAGATT 3718
Db 405 AGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGGAATATATAGGACAGATT 346

QY 3719 AATATTGTTCTGTCATGCCCAACATCCCTTTCTAAGGAAGACTGCCCTACTATAGCAG 3778
Db 345 AATATTGTTCTGTCATGCCCAACATCCCTTTCTAAGGAAGACTGCCCTACTATAGCAG 286

QY 3779 TTTTATATTGTCATTTATGAAATGAGGAGTCTCTGTTTAAACCMTTCCCAATTCGGGTGTGT 3838
Db 285 TTTTATATTGTCATTTATGAAATGAGGAGTCTCTGTTTAAACCMTTCCCAATTCGGGTGTGT 227

QY 3839 ACAATATTGGGTGTGTCAGTATTTTCCCTTTTAAACCMTTCCCAATTCGGGTGTGT 3898
Db 226 ACAATATTGGGTGTGTCAGTATTTTCCCTTTTAAACCMTTCCCAATTCGGGTGTGT 171

QY 3899 AGTGGATGTTTCCATTTGGGTGTTTAAATTTGTATATCCCTGATAGCTATTAATTTGGGTGAT 3958
Db 170 TGATG-----TTTCATTTTGGGTGTTTAAATTTGTATATCCCTGATAGCTATTAATTTGGGTGAT 116

QY 3959 AGAAATCTTTATACATTTAGATGCAAGTCTCTTGYCGGATATAGTATTGAGATATTA 4018
Db 115 AGAAATCTTTATACATTTAGATGCAAGTCTCTTGYCGGATATAGTATTGAGATATTA 56

QY 4019 CACCTAGTCTGGGTGATGCTGTTTCTTTATGCTCTTTTATGATGAATAGAGTTTT 4073
Db 55 CACCTAGTCTGGGTGATGCTGTTTCTTTATGCTCTTTTATGATGAATAGAGTTTT 1

RESULT 4
US-10-025-380-343/c
; Sequence 343, Application US/10025380
; Publication No. US2002018219A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C14
```

```
; CURRENT APPLICATION NUMBER: US/10/025,380
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 1129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 343
; LENGTH: 646
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-025-380-343

Query Match      12.5%; Score 552.2; DB 13; Length 646;
Best Local Similarity 93.0%; Pred. No. 5.8e-111;
Matches 609; Conservative 4; Mismatches 32; Indels 10; Gaps 3;

QY 3419 AAAAGCTTCTCAGCCTTCTAGGGAACAGAAATTTGGGTGAGCCCAATCTGCAATTTCTACT 3478
Db 645 AAAGCTTCTCAGCCTTCTAGGGAACAGAAATTTGGGTGAGCCCAATCTGCAATTTCTACT 586

QY 3479 ACAGGCATTGAGACCAAGTTAGATTATTGAATATTATAGAGAGTTATGAACACTTAATTT 3538
Db 585 ACAGGCATTGAGACCAAGTTAGATTATTGAATATTATAGAGAGTTATGAACACTTAATTT 526

QY 3539 ATGATAGTGTATGACATTCGATAGAACATGGGATCTTTAGAAAGTAGAATTCACAGGGC 3598
Db 525 ATGATAGTGTATGACATTCGATAGAACATGGGATCTTTAGAAAGTAGAATTCACAGGGC 466

QY 3599 ATATTAGTTGATGAAATGGAGTCATTTGAGTCTCTTTTAATAGCCATGTATCATTAATTTACCA 3658
Db 465 ATATTAGTTGATGAAATGGAGTCATTTGAGTCTCTTTTAATAGCCATGTATCATTAATTTACCA 406

QY 3659 AGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGGAATATATAGGACAGATT 3718
Db 405 AGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGGAATATATAGGACAGATT 346

QY 3719 AATATTGTTCTGTCATGCCCAACATCCCTTTCTAAGGAAGACTGCCCTACTATAGCAG 3778
Db 345 AATATTGTTCTGTCATGCCCAACATCCCTTTCTAAGGAAGACTGCCCTACTATAGCAG 286

QY 3779 TTTTATATTGTCATTTATGAAATGAGGAGTCTCTGTTTAAACCMTTCCCAATTCGGGTGTGT 3838
Db 285 TTTTATATTGTCATTTATGAAATGAGGAGTCTCTGTTTAAACCMTTCCCAATTCGGGTGTGT 227

QY 3839 ACAATATTGGGTGTGTCAGTATTTTCCCTTTTAAACCMTTCCCAATTCGGGTGTGT 3898
Db 226 ACAATATTGGGTGTGTCAGTATTTTCCCTTTTAAACCMTTCCCAATTCGGGTGTGT 171

QY 3899 AGTGGATGTTTCCATTTGGGTGTTTAAATTTGTATATCCCTGATAGCTATTAATTTGGGTGAT 3958
Db 170 TGATG-----TTTCATTTTGGGTGTTTAAATTTGTATATCCCTGATAGCTATTAATTTGGGTGAT 116

QY 3959 AGAAATCTTTATACATTTAGATGCAAGTCTCTTGYCGGATATAGTATTGAGATATTA 4018
Db 115 AGAAATCTTTATACATTTAGATGCAAGTCTCTTGYCGGATATAGTATTGAGATATTA 56

QY 4019 CACCTAGTCTGGGTGATGCTGTTTCTTTATGCTCTTTTATGATGAATAGAGTTTT 4073
Db 55 CACCTAGTCTGGGTGATGCTGTTTCTTTATGCTCTTTTATGATGAATAGAGTTTT 1

RESULT 5
US-09-969-034-1463/c
; Sequence 1463, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Astle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Poorima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
```











QY 3419 AAAAGCTTCTCAGCCTTCTAGGAAACAGAAATTGGGTGAGCCAACTCTCAATTTCTACT 3478  
DB 1 AAAAGCTTCTCAGCCTTCTAGGAAACAGAAATTGGGTGAGCCAACTCTCAATTTCTACT 60  
QY 3479 ACAGGCATTGAGACCAAGTTAGATTATTGAAATATATATAGAGAGTTATGAACACTTAAATT 3538  
DB 61 ACAGGCATTGAGACCAAGTTAGATTATTGAAATATATATAGAGAGTTATGAACACTTAAATT 120  
QY 3539 ATGATAGTGGTATGACATTTGGATAGAACATGGGATCTTTAGAGTAGAATTGACAGGCG 3598  
DB 121 ATGATAGTGGTATGACATTTGGATAGAACATGGGATCTTTAGAGTAGAATTGACAGGCG 180  
QY 3599 ATATTAGTTGATGAAATGAGTCAATTTGAGTCTTTTAAATAGCCATCTATCAATAATTACCA 3658  
DB 181 ATATTAGTTGATGAAATGAGTCAATTTGAGTCTTTTAAATAGCCATCTATCAATAATTACCA 240  
QY 3659 AGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTTAAAGNATATAATGACAGATT 3718  
DB 241 AGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTTAAAGNATATAATGACAGATT 300  
QY 3719 AATATTGTTTCTGTGATGCCACAAATCCCTTCTTAAGGAAGACTGCCCTACTATAGCAG 3778  
DB 301 AATATTGTTTCTGTGATGCCACAAATCCCTTCTTAAGGAAGACTGCCCTACTATAGCAG 360  
QY 3779 TTTTATATTGTCATTTATGAATATAATGAATGAGGAGTTCTGTACTCTCTGTCTTT 3838  
DB 361 TTTTATATTGTCATTTATGAATATAATGAATGAGGAGTTCTGTACTCTCTGTCTTT 419  
QY 3839 ACAAAATTTGGTGTGTCAGGATTTTCCCTTTTAAACMMTCCCAATTCGGGTGTT 3898  
DB 420 ACAAAATTTGGTGTGTCAGGATTTTCCCTTTTAAACMMTCCCAATTCGGGTGTT 475  
QY 3899 AGGTGATGTTCCATTTGGGTTTAAATTTGATATCCCTGATAGCTAATTTGGGTAT 3958  
DB 476 TGAATGTTTGGTTTAAATTTGATATCCCTGATAGCTAATTTGGGTAT 530  
QY 3959 AGAAATCTTTATACATTTAGATGCAAGTCTCTGCGGATATAGCTATTGAGATA 4015  
DB 531 AGAAATCTTTATACATTTAGATGCAAGTCTCTGCGGATATAGCTATTGAGATA 587

RESULT 10  
US-10-242-535A-26197  
; Sequence 26197, Application US/10242535A  
; Publication No. US20040013663A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2005  
; CURRENT APPLICATION NUMBER: US/10/242,535A  
; CURRENT FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 10/085,783  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 26197  
; LENGTH: 481  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (31)..(31)  
; OTHER INFORMATION: n is a, c, g, or t  
US-10-242-535A-26197  
Query Match 10.9%; Score 478.4; DB 17; Length 481;

Best Local Similarity 99.6%; Pred. No. 8.7e-95;  
Matches 479; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 3126 ATGTACCTGTGATCCAGCAAGAGGAGTTCAGTCAAGAGTCACTACAAGTATTAGTT 3185  
DB 1 ATGTACCTGTGATCCAGCAAGAGTTCAGTCAAGAGTCACTACAAGTATTAGTT 60  
QY 3186 GTTTAGAGAAATGAGAAATGGAACAGTGAAGGATGAGGCCATATTTCATGATTCCTT 3245  
DB 61 GTTTAGAGAAATGAGAAATGGAACAGTGAAGGATGAGGCCATATTTCATGATTCCTT 120  
QY 3246 GTTAAACAGAAACAGAAAGGACAAAGAGGCTTCTACATCACTCTCACCTTCCAAA 3305  
DB 121 GTTAAACAGAAACAGAAAGGACAAAGAGGCTTCTACATCACTCTCACCTTCCAAA 180  
QY 3306 TCTTGTGAAGTGCATCTACTTCCAGAACCAAAATTAATTAATTTCCAGTTCTGGCTGC 3365  
DB 181 TCTTGTGAAGTGCATCTACTTCCAGAACCAAAATTAATTAATTTCCAGTTCTGGCTGC 240  
QY 3366 TTGCAAGTGGAACTCCAGCTCAAGGGAGTTAGGGAATGAAGTCTTTTTTAAAGCT 3425  
DB 241 TTGCAAGTGGAACTCCAGCTCAAGGGAGTTAGGGAATGAAGTCTTTTTTAAAGCT 300  
QY 3426 TCTCAGCCTTCTCAGGAAACAGAAATTCGGTGAGCCAAATCTGCAATTTCTACTACAGCA 3485  
DB 301 TCTCAGCCTTCTCAGGAAACAGAAATTCGGTGAGCCAAATCTGCAATTTCTACTACAGCA 360  
QY 3486 TTGACACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3545  
DB 361 TTGACACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 420  
QY 3546 TGGTATGACATTTGATAGAACATGCGTACTTTTGAAGTAGAATTTGACAGGCGCATATTAG 3605  
DB 421 TGGTATGACATTTGATAGAACATGCGTACTTTTGAAGTAGAATTTGACAGGCGCATATTAG 480  
QY 3606 T 3606  
DB 481 T 481

RESULT 11  
US-10-085-783A-26197  
; Sequence 26197, Application US/10085783A  
; Publication No. US20040037841A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2002  
; CURRENT APPLICATION NUMBER: US/10/085,783A  
; CURRENT FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 26197  
; LENGTH: 481  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (31)..(31)  
; OTHER INFORMATION: n is a, c, g, or t  
US-10-085-783A-26197

Query Match 10.9%; Score 478.4; DB 18; Length 481;  
Best Local Similarity 99.6%; Pred. No. 8.7e-95;  
Matches 479; Conservative 0; Mismatches 2; Indels 0; Gaps 0;



QY 3126 ATGTACCTGTGATCCAGCAAGAGGAGTTCCAGTCAAGAGTCATCAACATGATTAGTT 3185  
Db 1 ATGTACCTGTGATCCAGCAAGAGTTCAGGAGTTTCAGTCAAGAGTCATCAACATGATTAGTT 60  
QY 3186 GTTTAGAGATGAGAAATGGAACAGTCAGGAATGGAGGCCATATTTCCATGACTTCCCTT 3245  
Db 61 GTTTAGAGATGAGAAATGGAACAGTCAGGAATGGAGGCCATATTTCCATGACTTCCCTT 120  
QY 3246 GTAACAGAAAGCAACAGAAAGGCAAGAGGCTGGCCCTCTACATCACTCTCACCTTCCAAA 3305  
Db 121 GTAACAGAAAGCAACAGAAAGGCAAGAGGCTGGCCCTCTACATCACTCTCACCTTCCAAA 180  
QY 3306 TCTTTGGAGATGATCTACTTCCAGAACCAATTAACCTTACTTCCAAAGTTCTGGCTGC 3365  
Db 181 TCTTTGGAGATGATCTACTTCCAGAACCAATTAACCTTACTTCCAAAGTTCTGGCTGC 240  
QY 3366 TTCAGGTGAACTCCAGCTGCAAGGAGTTAGGAAATCAAGTCTTTTTTAAAGCT 3425  
Db 241 TTCAGGTGAACTCCAGCTGCAAGGAGTTAGGAAATCAAGTCTTTTTTAAAGCT 300  
QY 3426 TCTCAGCCTTCTAGGGAACAGAAATTTGGGTGAGCCAACTCGCAATTTCTACTACAGGCA 3485  
Db 301 TCTCAGCCTTCTAGGGAACAGAAATTTGGGTGAGCCAACTCGCAATTTCTACTACAGGCA 360  
QY 3486 TTCAGACCAGTTAGATTATTTGAAATATTATAGAGAGTTATGAACACTTAAATTTATGATAG 3545  
Db 361 TTCAGACCAGTTAGATTATTTGAAATATTATAGAGAGTTATGAACACTTAAATTTATGATAG 420  
QY 3546 TGGTATGACATTCGATAGACATGGGATACCTTTAGAGTAGAATTCAGAGGCGATATTAG 3605  
Db 421 TGGTATGACATTCGATAGACATGGGATACCTTTAGAGTAGAATTCAGAGGCGATATTAG 480  
QY 3606 T 3606  
Db 481 T 481

RESULT 12  
US-10-242-535A-7179  
; Sequence 7179, Application US/10242535A  
; Publication No. US20040013663A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2005  
; CURRENT APPLICATION NUMBER: US/10/242,535A  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 10/085,783  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 7179  
; LENGTH: 439  
; TYPE: DNA  
; ORGANISM: Human  
US-10-242-535A-7179  
  
Query Match 10.0%; Score 438.2; DB 17; Length 439;  
Best Local Similarity 99.8%; Pred. No. 5.9e-86;  
Matches 437; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
QY 3344 CTTACTTCCAAAGTTCTGGCTGCTTGCAGGTGGAACTCCAGCTCGAAGGGAGTTAGGAAA 3403  
Db 1 CTTACTTCCAAAGTTCTGGCTGCTTGCAGGTGGAACTCCAGCTCGAAGGGAGTTAGGAAA 60  
QY 3404 TGAAGGTCTTTTTTAAAGCTTCTCAGCCTTCTCAGGGAACAGAAATTTGGGTGAGCCAA 3463  
Db 61 TGAAGGTCTTTTTTAAAGCTTCTCAGCCTTCTCAGGGAACAGAAATTTGGGTGAGCCAA 120  
QY 3464 TCTGCAATTTCTACTACAGGCATTGAGACCAAGTTAGATTATTTGAAATATATATAGAGATT 3523  
Db 121 TCTGCAATTTCTACTACAGGCATTGAGACCAAGTTAGATTATTTGAAATATATATAGAGATT 180  
QY 3524 ATGAACACTTAAATTTATGATAGTGGTATGACATTTGGATAGAACATCGGATACCTTTAGAAG 3583  
Db 181 ATGAACACTTAAATTTATGATAGTGGTATGACATTTGGATAGAACATCGGATACCTTTAGAAG 240

Db 61 TGAAGGTCTTTTTTAAAGCTTCTCAGCCTTCTCAGGGAACAGAAATTTGGGTGAGCCAA 120  
QY 3464 TCTGCAATTTCTACTACAGGCATTGAGACCAAGTTAGATTATTTGAAATATATATAGAGATT 3523  
Db 121 TCTGCAATTTCTACTACAGGCATTGAGACCAAGTTAGATTATTTGAAATATATATAGAGATT 180  
QY 3524 ATGAACACTTAAATTTATGATAGTGGTATGACATTTGGATAGAACATCGGATACCTTTAGAAG 3583  
Db 181 ATGAACACTTAAATTTATGATAGTGGTATGACATTTGGATAGAACATCGGATACCTTTAGAAG 240  
QY 3584 TAGAATTCAGAGGCATATTTAGTTGATGAAATGGAGTCAATTTGAGTCTTTTAATAGCCAT 3643  
Db 241 TAGAATTCAGAGGCATATTTAGTTGATGAAATGGAGTCAATTTGAGTCTTTTAATAGCCAT 300  
QY 3644 GTATCATATATTCACCAAGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGGAA 3703  
Db 301 GTATCATATATTCACCAAGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGGAA 360  
QY 3704 TATAATGGACAGATTAAATTTGTTTCTGTCTGTCGCCCAATCCCTTTCTAAGGAAGACT 3763  
Db 361 TATAATGGACAGATTAAATTTGTTTCTGTCTGTCGCCCAATCCCTTTCTAAGGAAGACT 420  
QY 3764 GCCTACTATAGCAGTTT 3782  
Db 421 GCCTACTATAGCAGTTT 439

RESULT 13  
US-10-085-783A-7179  
; Sequence 7179, Application US/10085783A  
; Publication No. US20040037841A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2002  
; CURRENT APPLICATION NUMBER: US/10/085,783A  
; CURRENT FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 7179  
; LENGTH: 439  
; TYPE: DNA  
; ORGANISM: Human  
US-10-085-783A-7179

Query Match 10.0%; Score 438.2; DB 18; Length 439;  
Best Local Similarity 99.5%; Pred. No. 5.9e-86;  
Matches 437; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
QY 3344 CTTACTTCCAAAGTTCTGGCTGCTTGCAGGTGGAACTCCAGCTCGAAGGGAGTTAGGAAA 3403  
Db 1 CTTACTTCCAAAGTTCTGGCTGCTTGCAGGTGGAACTCCAGCTCGAAGGGAGTTAGGAAA 60  
QY 3404 TGAAGGTCTTTTTTAAAGCTTCTCAGCCTTCTCAGGGAACAGAAATTTGGGTGAGCCAA 3463  
Db 61 TGAAGGTCTTTTTTAAAGCTTCTCAGCCTTCTCAGGGAACAGAAATTTGGGTGAGCCAA 120  
QY 3464 TCTGCAATTTCTACTACAGGCATTGAGACCAAGTTAGATTATTTGAAATATATATAGAGATT 3523  
Db 121 TCTGCAATTTCTACTACAGGCATTGAGACCAAGTTAGATTATTTGAAATATATATAGAGATT 180  
QY 3524 ATGAACACTTAAATTTATGATAGTGGTATGACATTTGGATAGAACATCGGATACCTTTAGAAG 3583  
Db 181 ATGAACACTTAAATTTATGATAGTGGTATGACATTTGGATAGAACATCGGATACCTTTAGAAG 240



QY 3584 TAGAATTGACGGGCATATTAGTTGATGAAATGGAGTCATTTGAGTCTVTTAATAGCCAT 3643  
Db |||||  
QY 241 TAGAATTGACGGGCATATTAGTTGATGAAATGGAGTCATTTGAGTCTCTTAATAGCCAT 300  
Db |||||  
QY 3644 GTATCATTAATTACCAAGTGAAGCTGCTGGAACATATGGTCTCCATTTTACAGTTAAGGAA 3703  
Db |||||  
QY 301 GTATCATTAATTACCAAGTGAAGCTGCTGGAACATATGGTCTCCATTTTACAGTTAAGGAA 360  
Db |||||  
QY 3704 TATAATGGACAGATTAATATGTTTCTGTCTATGCCCAAAATCCCTTCTTAAGGAAGACT 3763  
Db |||||  
QY 361 TATAATGGACAGATTAATATGTTTCTGTCTATGCCCAAAATCCCTTCTTAAGGAAGACT 420  
Db |||||  
QY 3764 GGCCTACTATAGCAGTTT 3782  
Db |||||  
QY 421 GGCCTACTATAGCAGTTT 439  
Db |||||  
RESULT 14  
US-09-918-995-13834  
; Sequence 13834, Application US/09918995  
; Publication No. US20030073623A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; FILE REFERENCE: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918,995  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/235,076  
; PRIOR FILING DATE: 1999-01-20  
; NUMBER OF SEQ ID NOS: 38054  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 13834  
; LENGTH: 468  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(468)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-918-995-13834  
Query Match 9.2%; Score 405.6; DB 10; Length 468;  
Best Local Similarity 97.4%; Pred. No. 9.2e-79;  
Matches 411; Conservative 1; Mismatches 10; Indels 0; Gaps 0;  
QY 2938 TGTCTTACTGAGCTGGGTGGTCTGTTTGGAGCTTATTAGAGTCTCTAGTTTTCCTA 2997  
Db |||||  
QY 47 TGTCTTACTGAGCTGGGTGGTCTGTTTGGAGCTTATTAGAGTCTCTAGTTTTCCTA 106  
Db |||||  
QY 2998 CTTTATAAGTAGAAATGGTGAGATTGTTTCTTTTCTTACCTAAAGGGAGATGGTAAGA 3057  
Db |||||  
QY 107 CTTTATAAGCAGAAACGGGAGATTGTTTGTGTTTCTTACCTTAAAGGGAGATGGCAAGA 166  
Db |||||  
QY 3058 ACAAATGAATGCTCTTTTCAAACTTATTGCAAGTGATTTTCAAGTCTGTGTTCAAAA 3117  
Db |||||  
QY 167 AACATGAATGCTCTTTTCAAACTTATTGACGAGTGATTTTACGCTGTGTTCAAAA 226  
Db |||||  
QY 3118 ATATATTTCATGTACCTGTGATCCAGCAAGAGGGAGTTCAGTCAAGAGTCACTACAAC 3177  
Db |||||  
QY 227 ATATATTTCATGTACCTGTGATCCAGCAAGAGGGAGTTCAGTCAAGAGTCACTACAAC 286  
Db |||||  
QY 3178 GATTAGTTGTTTAGAATGAGAAATGGAACAGTGAAGGATGGAGGCCATATTTCCATGA 3237  
Db |||||  
QY 287 GATTAGTTGTTTAAAGATGAGAAATGGAACAGTGAAGGATGGAGGCCATATTTCCATGA 346  
Db |||||  
QY 3238 CTTCCCTTGTAAACAGAACCAAGAGGACAGAGGCTGGCTCTTACATCACTCTCAC 3297  
Db |||||  
QY 347 CTTCCCTTGTAAACAGAACCAAGAGGACAGAGGCTGGCTCTTACATCACTCTCAC 406  
Db |||||  
QY 3298 CTTCCAAATCTTGTGGAAGTGCA TCTACTTGGCAGAACCAAAATTAATCTTACTTCCAGTT 3357  
Db |||||  
QY 407 CTTCCAAATCTTGTGGAAGTGCA TCTACTTGGCAGAACCAAAATTAATCTTACTTCCAGTT 466  
Db |||||

QY 3358 CT 3359  
Db |||||  
QY 467 CT 468  
Db |||||  
RESULT 15  
US-09-969-034-1974/c  
; Sequence 1974, Application US/09969034  
; Publication No. US20040110668A1  
; GENERAL INFORMATION:  
; APPLICANT: Burgess, Christopher C.  
; APPLICANT: Astle, Jon H.  
; APPLICANT: Carroll, Eddie III  
; APPLICANT: Catino, Theodore J.  
; APPLICANT: Dwivedi, Poornima  
; APPLICANT: Molino, Gary A.  
; APPLICANT: Thiagalingam, Arunthathi  
; APPLICANT: Lewis, Marcia E.  
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially  
; FILE REFERENCE: 1657/1032  
; CURRENT APPLICATION NUMBER: US/09/969,034  
; CURRENT FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: 60/237,271  
; PRIOR FILING DATE: 2000-02-10  
; NUMBER OF SEQ ID NOS: 4494  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1974  
; LENGTH: 483  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 36, 41, 58, 70, 85, 118, 128, 133, 148, 152, 183, 187, 195,  
; LOCATION: 203, 211, 217, 247, 289, 301, 303, 315, 324, 331, 342, 346,  
; LOCATION: 358, 367, 378, 402, 405, 412, 417, 420, 427, 429, 432, 442,  
; LOCATION: 444, 470, 475, 477  
; OTHER INFORMATION: n = A,T,C or G  
US-09-969-034-1974

Query Match 8.2%; Score 362; DB 11; Length 483;  
Best Local Similarity 83.8%; Pred. No. 3.8e-69;  
Matches 383; Conservative 0; Mismatches 74; Indels 0; Gaps 0;  
QY 1253 CCCACTGTAAAGAGTCTCTGCAAAACAAGCTCTAATAATAAGTAGTATGTTATCAATACT 1312  
Db |||||  
QY 473 CCCNCTGTAAAGAGTCTCTGCAAAACAAGTNTAATAAANAGNANGGTATCNAANACT 414  
Db |||||  
QY 1313 TTGGCTTAAGATCAGATCCAAACTATCAGCTTTCCACCAACTAAATTTGCCAAGTATAAAT 1372  
Db |||||  
QY 413 TTGGTTAANATNAGATCCAAACTATCAGCTTTCCNCAATTAANTGGCCAGNTATA 354  
Db |||||  
QY 1373 AAAAGTAAAGATAGGGCTTTCTCAGCAGCAGCAGACCAACTCCATCAGAAACTACTTTTCAG 1432  
Db |||||  
QY 353 ATAAAGGNAAGTGGTTTTTCANGCAGCANCAGCAANTCCATCAGAAANTNCTTTTCAG 294  
Db |||||  
QY 1433 CCGTCTACCAAAAAAGGAAAGGATGAAGAAATCAAGAAATGTCTTCATGCAATCA 1492  
Db |||||  
QY 293 CCGTNTACCAAAAAAGGAAAGGATGAAGAAATCAAGAAATGTNTTTCATGCAATCA 234  
Db |||||  
QY 1493 GCAAGATGAAGACCTGTGTTCTCTTTTAGAAACAAACCAACCTGCTACACCTCATTTG 1552  
Db |||||  
QY 233 GCAAGATGAAGACCTGTGTTCTTTTANAAACAAACCAACCTGNTACNCCCTCATTTG 174  
Db |||||  
QY 1553 TGGAAAAATTAAGAGCAGCATCTATCTGAGAAATGAGCCTGTGGACAAAACCTCAGCAAT 1612  
Db |||||  
QY 173 TGGAAAAATTAAGAGCAGCATNTATNTGAGATGAGCCTGNGGACNCAAACTCAGNCAAT 114  
Db |||||  
QY 1613 AACTTATTTACAGATACAGATTTAAATCTATTGTGAAAAATTTGCCAGTAAATCTCAT 1672  
Db |||||  
QY 113 AACTTATTTACAGATACAGATTTAAATNTATTGTGAAAAATTTTGGCCAGTAAATNTCAT 54  
Db |||||



Qy 1673 GCTGCAGAAAAGCTTAAGATCAAAATAAAAAAGGGAAA 1709  
|||||  
Db 53 GCTGCAGAAAAGNTAAGNTCCAAAAA 17  
|||||

Search completed: August 30, 2005, 15:27:50  
Job time : 2583 secs



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2005, 05:15:28 ; Search time 13109 Seconds  
(without alignments)  
12784.879 Million cell updates/sec

Title: US-09-837-138-1  
Perfect score: 4403  
Sequence: 1 ttccggcacaggcgcgggttg.....accgcggtgagctccagct 4403

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_ges1:\*  
9: gb\_ges2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	4053.6	92.1	4692	3	HSMB06925
2	4030.2	91.5	4550	3	BC040519 Homo sapi
3	1460.2	33.2	1497	3	BC005293 Homo sapi
4	1446.8	32.9	2265	9	AY416034 Homo sapi
5	1427	32.4	1460	3	BC016762 Homo sapi
6	1090.4	24.8	1792	3	BC031900 Mus muscu
7	1083	24.6	1731	3	AK031933 Mus muscu
8	913.2	20.7	964	5	EX405940 BX405940
9	844.2	19.2	1021	5	EX405939 BX405939
10	834.2	18.9	1163	4	BM542698 AGENCOURT
11	822.8	18.7	927	4	BG182890 RST1768 A
12	822.2	18.7	1803	9	AY416036 Mus muscu
13	771.4	17.5	785	7	CN304420 170006000
14	741.4	16.8	784	6	CB250418 UI-CF-FNO
15	741	16.8	789	4	BG214621 RST34264
16	734.8	16.7	747	7	CN304425 170004239
17	732.8	16.6	742	7	CN304424 170005999
18	732.6	16.6	1036	5	BX422568 BX422568
19	732.4	16.6	902	4	BM461758 AGENCOURT
20	731.8	16.6	932	5	BUI16634 AGENCOURT
21	725.4	16.5	775	4	BG67675 602623363
22	717.4	16.3	2285	9	AY416035 Pan trogl
23	715.8	16.3	771	5	BUE20472 UI-H-FL1-
24	708.6	16.1	870	5	BUI177917 AGENCOURT

25	708.2	16.1	818	7	CF593314
26	706.6	16.0	763	4	BM014420
27	702.8	16.0	855	5	BU517247
28	702.4	16.0	811	5	BQ574642 UI-H-EZ1-
29	696.2	15.8	960	4	BG284646
30	690.8	15.7	779	4	BG388866
31	689.4	15.7	691	6	CB123692 K-EST0172
32	687.8	15.6	712	1	AI796269 wh44910.x
33	676.2	15.4	735	4	BG779728 AGENCOURT
34	674	15.3	779	5	BU929703 AGENCOURT
35	673.2	15.3	707	5	BU661996 cl79f11.z
36	672.6	15.3	807	4	RG197194 RST16433
37	671	15.2	752	1	AUI18357 AUI18357
38	669.2	15.2	718	2	AW976050 EST388159
39	664.4	15.1	747	6	CA777874
40	663	15.1	796	5	BM978914 UI-CF-DU1
41	663	15.1	845	4	BG532783
42	660.6	15.0	691	2	AW978306 EST390415
43	660.6	15.0	751	7	CN304418 170006001
44	659.8	15.0	917	5	BP381003 BP381003
45	651.2	14.8	706	4	BG292394 602386511

ALIGNMENTS

RESULT 1  
LOCUS HSM806925 4692 bp mRNA linear HTC 22-SEP-2004  
DEFINITION Homo sapiens mRNA; cDNA DKFP686G19151 (from clone DKFP686G19151).  
ACCESSION BX640816  
VERSION BX640816.1 GI:34365074  
KEYWORDS HTC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 4692)  
AUTHORS Bloembergen, H., Boecher, M., Brandt, P., Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.  
CONSTRM The German cDNA Consortium  
TITLE Direct Submission  
JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY  
COMMENT Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.  
This clone (DKFP686G19151) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering:  
http://www.rzpd.de/cgi-bin/products/ci.cgi?CloneID=DKFP686G19151  
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.  
FEATURES  
Location/Qualifiers  
1..4692  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="RZPD:DKFP686G19151Q"  
/db\_xref="taxon:9606"  
/clone="DKFP686G19151"  
/tissue type="colon endothel, primary cell culture"  
/clone\_libs="686 (synonym: hicc3). Vector pSport1\_Sfi; host DH10B; sites SfiIA + SfiIB"  
/dev\_stage="adult"  
/note="Nijmegen breakage syndrome 1 (nibrin), differentially spliced, alternative start"  
join(407..1555,1557..2426)  
/gene="DKFP686G19151"  
join(407..1555,1557..2426)  
/gene="DKFP686G19151"  
/codon\_start=1



/product="hypothetical protein"  
 /protein\_id="CAH56160.1"  
 /db\_xref="GI:52545912"  
 /translation="MQNGFSRTLKSGDGIITFGVFGSKFRIEYELIVACSSCLDVSGKT  
 ALNQAILQGGFTVNNWTECHLVMVSVKVTIKTICALICGRPIVKPEYFTEFLKAV  
 ESKQPPQIETSPVPLDEPSIGSNVDLSGRERKQIFPKGTFFILNKHKLKLSAV  
 VFGGGEARLITENEBEHFFLAGPVCVVDGTGTNSOTLI PDCKKWIQS IMDMLQRO  
 GLRP IPEAIGLAVIPMTKNYCDPOGHPSTGLKTTTPGPSLSQGVSDKLMPSAPV  
 NNTYVADTESQAUTWDLSERPKEIKVSKMEQKFRMLSDQAPTVKESCKTSSNNNSM  
 VNTLIAMRIPNYQLSPITKLPSINKSRASQQQNTSIRNYFPQSTPKRKEENDEQE  
 MESSKARLETSCSLBOPTAPSLMNKKEOHLSENEPVDTSNNLFTDRLDSIV  
 KNSAKSHAAEKLRSNKRKEMDDVAIEDEVLQKFDKTKPELEIDVKVQBEDVNSR  
 KPRMDIETNDTPSDEAVESSKI SOENIEIGKRRELKEDLSAKBISNNDKLODVS  
 MLPKLLLETFPSLVIKNSTSRNPSGINDDYQLKNFKFKVTVPGAKLPHIIGGS  
 DLIAHARXNTELEWLRQEMEVQNHAKBESLADDLFRYNFLKRRR"

## ORIGIN

Query Match 92.1%; Score 4053.6; DB 3; Length 4692;  
 Best Local Similarity 96.0%; Pred. No. 0;  
 Matches 4287; Conservative 9; Mismatches 28; Indels 142; Gaps 8;

QY 13 CGCGGTTGCA CGTTCGCCGCCAGCCCTGAGAGCCGACCGATGTGGAAACTGCTGCCCGC 72  
 DB 71 CGCGGTTGCA CGTTCGCCGCCAGCCCTGAGAGCCGACCGATGTGGAAACTGCTGCCCGC 130

QY 73 CGCGGCCGCGCAGGAGGAGAACCATACAGACTTTTGTACTGGGTTGAGTACGTTGTTGG 132  
 DB 131 CGCGGCCGCGCAGGAGGAGAACCATACAGACTTTTGTACTGGGTTGAGTACGTTGTTGG 190

QY 133 AAGGAAAAACTGTGCCATTTCTAATGAAAAATGATCGATCGATCGAGCCGAAATCATGCTGT 192  
 DB 191 AAGGAAAAACTGTGCCATTTCTGATGAAAAATGATCGATCGATCGAGCCGAAATCATGCTGT 250

QY 193 GTTAACTGCTAACTTTTCTGTAAACCAACCTG----- 223  
 DB 251 GTTAACTGCTAACTTTTCTGTAAACCAACCTGGTATAGAAAAACATTTTGGAAATTTCTTAC 310

QY 224 -----AGTCAACAGATGAATCCCTGCTATTGACATTAAGAT 262  
 DB 311 AGAAAAAGTAAATTTGTTACTAGTCAACAGATGAATCCCTGTTATGACATTAAGAT 370

QY 263 AATTCTAAGTATGGTACCTTTGTTAATGAGGAAAAAATGCAGAAATGGCTTTTCCCGAACT 322  
 DB 371 AATTCTAAGTATGGTACCTTTGTTAATGAGGAAAAAATGCAGAAATGGCTTTTCCCGAACT 430

QY 323 TTGAAGTCGGGGATGGTATTAATTTTGGAGTGTGTTGGAATGAAATTCAGAAATAGAGTAT 382  
 DB 431 TTGAAGTCGGGGATGGTATTAATTTTGGAGTGTGTTGGAATGAAATTCAGAAATAGAGTAT 490

QY 383 GAGCCTTTGGTGCATGCTCTCTCTGTTTAGATGCTCTGCGGAAACTGCTTTTAAATCAA 442  
 DB 491 GAGCCTTTGGTGCATGCTCTCTCTGTTTAGATGCTCTGCGGAAACTGCTTTTAAATCAA 550

QY 443 GCTATATTGCAACTTCGAGGATTTACTGTAAACAAATTGGACAGAAATGCCACTCACCTT 502  
 DB 551 GCTATATTGCAACTTCGAGGATTTACTGTAAACAAATTGGACAGAAATGCCACTCACCTT 610

QY 503 GTCATGGTATCAGTGAAGTTTACCATTAAACCAATATGTGCACTCAATTTGTGAACGTCCA 562  
 DB 611 GTCATGGTATCAGTGAAGTTTACCATTAAACCAATATGTGCACTCAATTTGTGAACGTCCA 670

QY 563 ATTGTAAGCCAGAAATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAAGAGAGCCCT 622  
 DB 671 ATTGTAAGCCAGAAATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAAGAGAGCCCT 730

QY 623 CCACAAATTTGAAAGTTTTTACCACCTCTTTGATGAACCATCTATTGGAAGTAAAAATGTT 682  
 DB 731 CCACAAATTTGAAAGTTTTTACCACCTCTTTGATGAACCATCTATTGGAAGTAAAAATGTT 790

QY 683 GATCTGTCAAGCGGAGGAAAGAAACAAATCTTCAAAGGGAAACATTTATATATTTTGG 742  
 DB 791 GATCTGTCAAGCGGAGGAAAGAAACAAATCTTCAAAGGGAAACATTTATATATTTTGG 850

743 AATGCCAAACAGCATAGAATAATTGATTTCCCGAGTTGTCTTTGGAGGTGGGGAAGCTAGG 802  
 DB 851 AATGCCAAACAGCATAGAATAATTGATTTCCCGAGTTGTCTTTGGAGGTGGGGAAGCTAGG 910

QY 803 TTGTAAACAGAGAGAAATGAAGAAGAAATTAATTTCTTTTGGTCTCCGGGAACGTGTGTT 862  
 DB 911 TTGTAAACAGAGAGAAATGAAGAAGAAATTAATTTCTTTTGGTCTCCGGGAACGTGTGTT 970

QY 863 GTTGATCAGAGAAATAACAACTCAGAGACCTTAATTCCTGACTGTGCAAGAAATGGATT 922  
 DB 971 GTTGATCAGAGAAATAACAACTCAGAGACCTTAATTTCTGACTGTGCAAGAAATGGATT 1030

QY 923 CAGTCAATAATGATATGCTCCAAAGGCAAGGTCTTAGACCTATTCTCTGAAGCAGAAAT 982  
 DB 1031 CAGTCAATAATGATATGCTCCAAAGGCAAGGTCTTAGACCTATTCTCTGAAGCAGAAAT 1090

QY 983 GGATTTGGCGGTGATTTTCATGACTACAAGAAATTAATCTGTGATCTCTCAGGGCCATCCCCAGT 1042  
 DB 1091 GGATTTGGCGGTGATTTTCATGACTACAAGAAATTAATCTGTGATCTCTCAGGGCCATCCCCAGT 1150

QY 1043 ACAGAGTTAAAGACAACTCCAGGACCAAGCCCTTTCACAGGCGGTGTCAGTTGATGAA 1102  
 DB 1151 ACAGAGTTAAAGACAACTCCAGGACCAAGCCCTTTCACAGGCGGTGTCAGTTGATGAA 1210

QY 1103 AAACCTAAATGCGCAAGCGCCCTCAGTGAACACTACAACTACGTAGCTGACACAGAAATCAGAG 1162  
 DB 1211 AAACCTAAATGCGCAAGCGCCCTCAGTGAACACTACAACTACGTAGCTGACACAGAAATCAGAG 1270

QY 1163 CAAGCAGATACATGGGATTTGAGTGAAGGCGCAAGAAATCAAAAGTCTCCAAATATGAA 1222  
 DB 1271 CAAGCAGATACATGGGATTTGAGTGAAGGCGCAAGAAATCAAAAGTCTCCAAATATGAA 1330

QY 1223 CAAAATTTCAGAAATGCTTTTCAAGACCGACCCACTGTAAAGGAGTCTGCAAAACAGC 1282  
 DB 1331 CAAAATTTCAGAAATGCTTTTCAAGACCGACCCACTGTAAAGGAGTCTGCAAAACAGC 1390

QY 1283 TCTAATAATAATAGTATGATGATCAAACTTTGGCTAAGATGAGAAATCCCAAACTATCAG 1342  
 DB 1391 TCTAATAATAATAGTATGATGATCAAACTTTGGCTAAGATGAGAAATCCCAAACTATCAG 1450

QY 1343 CTTTCCACCAACTAAATTCGCAAGTATAATAAAGTAAAGATAGGGCTTCTCAGCAGCAG 1402  
 DB 1451 CTTTCCACCAACTAAATTCGCAAGTATAATAAAGTAAAGATAGGGCTTCTCAGCAGCAG 1510

QY 1403 CAGACCACTCCATCAGAACTACTTTTCAGCGCTCTACC-AAAAAAGGGAAGGAGTGA 1461  
 DB 1511 CAGACCACTCCATCAGAACTACTTTTCAGCGCTCTACCAAAAAAGGGAAGGAGTGA 1570

QY 1462 AGAAATCAAGAAATGCTCTTCATGCAAAATCAGCAAGAAATAGAAACGTCTTGTCTCTTTT 1521  
 DB 1571 AGAAATCAAGAAATGCTCTTCATGCAAAATCAGCAAGAAATAGAAACGTCTTGTCTCTTTT 1630

QY 1522 AGAACAAACACAACTGCTTACACCTCATTTGTGAAAAAATAAGGAGCAGCATCTATCTGA 1581  
 DB 1631 AGAACAAACACAACTGCTTACACCTCATTTGTGAAAAAATAAGGAGCAGCATCTATCTGA 1690

QY 1582 GAATGAGCCTGTGGACACAACTCAGACAACTTAATTTTACAGATCAGATTTTAAATC 1641  
 DB 1691 GAATGAGCCTGTGGACACAACTCAGACAACTTAATTTTACAGATCAGATTTTAAATC 1750

QY 1642 TATTGTGAAAAATTCCTGCCAGTAAATCTCATGCTCAGAAAAAGCTAAAGTCAAAATAAAA 1701  
 DB 1751 TATTGTGAAAAATTCCTGCCAGTAAATCTCATGCTCAGAAAAAGCTAAAGTCAAAATAAAA 1810

QY 1702 AAGGAAATGAGATGATGTGGCCATAGAGATGAAGTATTTGGAACAGTATTTCAGAGCAC 1761  
 DB 1811 AAGGAAATGAGATGATGTGGCCATAGAGATGAAGTATTTGGAACAGTATTTCAGAGCAC 1870

QY 1762 AAAACCCAGTGTAGAAATTTGATGCAAGTTCAAAAACAGGAGGAGATGTCAATGTTAG 1821  
 DB 1871 AAAACCCAGTGTAGAAATTTGATGCAAGTTCAAAAACAGGAGGAGATGTCAATGTTAG 1930

QY 1822 AAAAGGCCAAGGATGGATATAGAAAACAATGACACTTTTCAGTGTAGGAGCAGTACCAGA 1881



	Db	1931	 AAAAAGCCCAAGGATGGATATAGAAAACAAATGCACCTTCAGTGATGAAGCAGTACCAGA	1990
	Qy	1882	AAGTAGCAAATAATCTCAAGAAAATGAAATTGGGAAGAAACGTGAACCTCAAGGAAGACTC	1941
	Db	1991	AAGTAGCAAATAATCTCAAGAAAATGAAATTGGGAAGAAACGTGAACCTCAAGGAAGACTC	2050
	Qy	1942	ACTATGTCAGCTAAGAATAATCTACAATAGCAAACTTCAGSAGATGATGAGATGCT	2001
	Db	2051	ACTATGTCAGCTAAGAATAATCTACAATAGCAAACTTCAGSAGATGATGAGATGCT	2110
	Qy	2002	TCCAAAAAGCGTTATTGTGACTGNAITTAGATCACTGCTGATTAATAAACCTCTACTTCCAG	2061
	Db	2111	TCCAAAAAGCGTTATTGTGACTGNAITTAGATCACTGCTGATTAATAAACCTCTACTTCCAG	2170
	Qy	2062	AAATCCCGCTGGCAATAATGATGATTATGTCCTCACTAAAAAATTTCAAGAAAATTCAAAA	2121
	Db	2171	AAATCCCATCTGGCATAATGATGATTATGTCCTCACTAAAAAATTTCAAGAAAATTCAAAA	2230
	Qy	2122	GGTCACATATCTCTGGAGCAGGAAAACTTCACACATCATTTGGAGGATCAGATCTAATAGC	2181
	Db	2231	GGTCACATATCTCTGGAGCAGGAAAACTTCACACATCATTTGGAGGATCAGATCTAATAGC	2290
	Qy	2182	TCATCATGCTCGAAAAGAAATACAGAACTTAGAAGAGTGGCTAAGCGAGGAAATGAGGTACA	2241
	Db	2291	TCATCATGCTCGAAAAGAAATACAGAACTTAGAAGAGTGGCTAAGCGAGGAAATGAGGTACA	2350
	Qy	2242	AAATCAACATGCAAAAGAAGAGTCTCTTGCTGTGATGATCTTTTTAGATACAAATCCTTATTT	2301
	Db	2351	AAATCAACATGCAAAAGAAGAGTCTCTTGCTGTGATGATCTTTTTAGATACAAATCCTTATTT	2410
	Qy	2302	AAAAAGGAGAGATAACTGAGGAGTTTTTAAAAAGAAGCCATGGAAAACTTCTCTAGTAAGC	2361
	Db	2411	AAAAAGGAGAGATAACTGAGGAGTTTTTAAAAAGAAGCCATGGAAAACTTCTCTAGTAAGC	2470
	Qy	2362	ATCTACTTCAGGCCAACAGGTTATATGAATATATAGTGATATAGAAGCGATTTAAGTTAC	2421
	Db	2471	ATCTACTTCAGGCCAACAGGTTATATGAATATATAGTGATATAGAAGCGATTTAAGTTAC	2530
	Qy	2422	AATGTTTTATGGCCTAAATTTATTAATAAAATGSCACAAAACCTTGATCTCTTTGTTATGT	2481
	Db	2531	AATGTTTTATGGCCTAAATTTATTAATAAAATGSCACAAAACCTTGATCTCTTTGTTATGT	2590
	Qy	2482	AACAAATGTTGTGTCGTTTTCAGGCTTTGTCATTTGCATCTTTTTTTTCATTTTTAAATGT	2541
	Db	2591	AACAAATGTTGTGTCGTTTTCAGGCTTTGTCATTTGCATCTTTTTTTTCATTTTTAAATGT	2650
	Qy	2542	GTTTTGTTTTATTAATAGTTAAATAGTCAAGTTCAAAATCTCTAAATRTACGTAAAGGTA	2601
	Db	2651	GTTTTGTTTTATTAATAGTTAAATAGTCAAGTTCAAAATCTCTAAATRTACGTAAAGGTA	2710
	Qy	2602	AAGGACTAAAGTCAACCTTCCACATTTGCTCCTAGCTACT-----	2640
	Db	2711	AA--GACTAAAGTCAACCTTCCACATTTGCTCCTAGCTACTTTTCCCTCAGAAAAAAT	2769
	Qy	2641	-----	2642
	Db	2770	TCATGATACATTTCTTTATGAATCTTTTCAGGGAATTTTGTAGTCCTATTCAAATTCCTA	2829
	Qy	2643	TTTTTAAATATTTCTCTACACAAATGATAGCAATAATATGCGAGTGTTCTACACCTTGCT	2702
	Db	2830	TTTTTAAATATTTCTCTACACAAATGATAGCAATAATATGCGAGTGTTCTACACCTTGCT	2889
	Qy	2703	TTTTTACTTAGTAAGATTAAANAATTAGGAATATCAATATATGTTTTTAAATATTTTTT	2762
	Db	2890	TTTTTACTTAGT-AGATTAAANAATTAGGAATATCAATATATGTTTTTAAATATTTTTT	2948
	Qy	2763	CTTTTCCATTTATGCTGTAGTCTTACCTTAAACTCTGGTGATGCCAAACAAAATGCGTTCAGT	2822
	Db	2949	CTTTTCCATTTATGCTGTAGTCTTACCTTAAACTCTGGTGATGCCAAACAAAATGCGTTCAGT	3008
	Qy	2823	GGTGCAAGTGCCTACATGTTATTCTAGTACTTAGAATACTGAAGACCATGTGCGAGACTT	2882







[illegible]







Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, J., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/INL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 16 Row: j Column: 6  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 3098674  
 This clone has the following problem: frame shifted.

## FEATURES

Location/Qualifiers  
 1. 1497  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3997534"  
 /tissue\_type="Bladder, carcinoma"  
 /clone\_lib="NIH\_MGC\_53"  
 /lab\_host="DH10B"  
 /note="vector: pDNR-LIB"

## ORIGIN

Query Match 33.2%; Score 1460.2; DB 3; Length 1497;  
 Best Local Similarity 99.4%; Pred. No. 4.9e-302;  
 Matches 1476; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 22 ACCTGGCCCGCCAGCCCTGAGAGCGGACCGATGTGGAAACTGCTGCCCGCCCGGGCCC 81  
 DB 2 ACCTGGCCCGCCAGCCCTGAGAGCGGACCGATGTGGAAACTGCTGCCCGCCCGGGCCC 61

QY 82 GGCAGGAGGAGACCATACAGACTTTTGACTGGCGTTGAGTACGTGTTGGAGGAAAA 141  
 DB 62 GGCAGGAGGAGACCATACAGACTTTTGACTGGCGTTGAGTACGTGTTGGAGGAAAA 121

QY 142 CTGTGCCATTCTAATTCGAAATGATCAGTCGATCAGCGGAAATCATGCTGTGTTAACTGC 201  
 DB 122 CTGTGCCATTCTAATTCGAAATGATCAGTCGATCAGCGGAAATCATGCTGTGTTAACTGC 181

QY 202 TAACTTTTCTGTAAACCACTGAGTCAACAGAGTAAATCCCTGTATTTGACATTAAGA 261  
 DB 182 TAACTTTTCTGTAAACCACTGAGTCAACAGAGTAAATCCCTGTATTTGACATTAAGA 241

QY 262 TAATTCAGTATGTTACCTTTGTTAATGAGGAAAAATGCAAGTGGCTTTTCCCGAAC 321  
 DB 242 TAATTCAGTATGTTACCTTTGTTAATGAGGAAAAATGCAAGTGGCTTTTCCCGAAC 301

QY 322 TTTGAAGTCGGGGATGGTATTACTTTTGGAGTGTGGAGTAAATTCAGAATAGAGTA 381  
 DB 302 TTTGAAGTCGGGGATGGTATTACTTTTGGAGTGTGGAGTAAATTCAGAATAGAGTA 361

QY 382 TGAGCCTTTGGTTGCAATGCTCTTTCTTTTGTAGATGCTCTGGGAAACTGCTTTAAATCA 441  
 DB 362 TGAGCCTTTGGTTGCAATGCTCTTTCTTTTGTAGATGCTCTGGGAAACTGCTTTAAATCA 421

QY 442 AGCTATATTCGAATCTGGAGGATTTACTGTAACAAATTTGACAGAGAAATGCACTACCT 501  
 DB 422 AGCTATATTCGAATCTGGAGGATTTACTGTAACAAATTTGACAGAGAAATGCACTACCT 481

QY 502 TGTTCATGGTATCAGTGAAGTTACCAATTAACAAATATGTGCACTCATTTGTGGAGTCC 561  
 DB 482 TGTTCATGGTATCAGTGAAGTTACCAATTAACAAATATGTGCACTCATTTGTGGAGTCC 541

QY 562 AATTGTAAGCCAGAAATATTTTACTGAATTCCTGAAAGCAGTTTCAGTCCAAAGAGCAGCC 621  
 DB 542 AATTGTAAGCCAGAAATATTTTACTGAATTCCTGAAAGCAGTTTCAGTCCAAAGAGCAGCC 601

QY 622 TCACAAATTTGAAGTTTACCACCTCTTGATGAACCAATCTATTGGAGTAAAAATGT 681  
 DB 602 TCACAAATTTGAAGTTTACCACCTCTTGATGAACCAATCTATTGGAGTAAAAATGT 661

QY 682 TGATCTGTGAGGAGCGAGGAGAAACAAATCTTCAAGGGAACATTTATATTTT 741  
 DB 662 TGATCTGTGAGGAGCGAGGAGAAACAAATCTTCAAGGGAACATTTATATTTT 721

QY 742 GAATGCCAAACAGCATAGAAATGTAGTTCCGAGTTGTCTTTGGAGGTGGGAAAGCTAG 801  
 DB 722 GAATGCCAAACAGCATAGAAATGTAGTTCCGAGTTGTCTTTGGAGGTGGGAAAGCTAG 781

QY 802 GTTGATTAACAGAGAGAAATGAGAAACATAATTTCTTTTGGCTCCGGGAAAGCTGTGT 861  
 DB 782 GTTGATTAACAGAGAGAAATGAGAAACATAATTTCTTTTGGCTCCGGGAAAGCTGTGT 841

QY 862 TGTGTATACAGGAATTAACAACTCAGAGACTTTAATTCCTGACTGTCTCAGAAAGAAATGGAT 921  
 DB 842 TGTGTATACAGGAATTAACAACTCAGAGACTTTAATTCCTGACTGTCTCAGAAAGAAATGGAT 901

QY 922 TCAGTCAATAATGGATATGCTCCAAAGGCAAGGTCTTAGACCTATTCTCTGAAGCAAAAT 981  
 DB 902 TCAGTCAATAATGGATATGCTCCAAAGGCAAGGTCTTAGACCTATTCTCTGAAGCAAAAT 961

QY 982 TGGATTGGCGGTGATTTTCATGACTACAAAGAAATTAATCTGTGACTCTCAGGGCCATCCCAG 1041  
 DB 962 TGGATTGGCGGTGATTTTCATGACTACAAAGAAATTAATCTGTGACTCTCAGGGCCATCCCAG 1021

QY 1042 TACAGGATTAAGAGCAACAACTCCAGGACCAAGCCCTTTTCAAGGGCGTGTCAAGTTGATGA 1101  
 DB 1022 TACAGGATTAAGAGCAACAACTCCAGGACCAAGCCCTTTTCAAGGGCGTGTCAAGTTGATGA 1081

QY 1102 AAAAATAATGCCAAGCGCCCGAGTGAACACTCAACATACATAGCTAGTACACAGAAATCAGA 1161  
 DB 1082 AAAAATAATGCCAAGCGCCCGAGTGAACACTCAACATACATAGCTAGTACACAGAAATCAGA 1141

QY 1162 GCAAGCAGATACATGGATTTGAGTGAAGGCGCAAGAAATCAAAAGTCTCCAAATGGA 1221  
 DB 1142 GCAAGCAGATACATGGATTTGAGTGAAGGCGCAAGAAATCAAAAGTCTCCAAATGGA 1201

QY 1222 ACAAAATTCAGAAATGCTTTTCAAGACGACCCACTGTAAAGGAGTCTCTGCAAAACAAG 1281  
 DB 1202 ACAAAATTCAGAAATGCTTTTCAAGACGACCCACTGTAAAGGAGTCTCTGCAAAACAAG 1261

QY 1282 CTCTAATAATAATAGTATGATCAAAATCTTTTGGCTAAGATGAGAAATCCAAATATATCA 1341  
 DB 1262 CTCTAATAATAATAGTATGATCAAAATCTTTTGGCTAAGATGAGAAATCCAAATATATCA 1321

QY 1342 GCTTTTCAACCACTTAATTTGCCAAGTATAATAAGTAAGATAGGCTTCTTCAGCAGCA 1401  
 DB 1322 GCTTTTCAACCACTTAATTTGCCAAGTATAATAAGTAAGATAGGCTTCTTCAGCAGCA 1381

QY 1402 GCAGACCAACTCCATCAGAAATCTCTTTCAGCGCTCTACC-AAAAAAGGGAAGGAGTG 1460  
 DB 1382 GCAGACCAACTCCATCAGAAATCTCTTTCAGCGCTCTACC-AAAAAAGGGAAGGAGTG 1441

QY 1461 AAGAAATCAAGAAATGTCTTCATGCAATTCAGCAAGATAGAAA 1505  
 DB 1442 AAGAAATCAAGAAATGTCTTCATGCAAAATAAAAAAAAAAAAAA 1486

RESULT 4  
 AY416034  
 LOCUS  
 DEFINITION Homo sapiens NBS1 gene, VIRTUAL TRANSCRIPT, partial sequence,  
 genomic survey sequence.  
 ACCESSION AY416034  
 VERSION AY416034.1 GI:39771994  
 KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,  
 Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, P., Murphy, B.,  
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sniensky, J.J.,  
 Adams, M.D. and Cargill, M.  
 Inferring nonneutral evolution from human-chimp-mouse orthologous  
 gene trios  
 JOURNAL Science 302 (5652), 1960-1963 (2003)







[illegible]

RESULT 5	BC016762	LOCUS	BC016762	1460 bp	mRNA	linear	HTC 09-NOV-2001
DEFINITION Homo sapiens, similar to Nijmegen breakage syndrome 1 (nibrin), clone IMAGE:4104186, mRNA.							

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnlnl.gov>  
Series: IRAL Plate: 32 Row: e Column: 20  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6996019  
This clone has the following problem: frame shifted

/db_xref="taxon:9606" /clone="IMAGE:4104186" /tissue_type="Brain, glioblastoma" /clone_lib="NIH_MGC_57" /lab_host="DH10B" /note="vector: pDNR-LIB"									
ORIGIN	Query Match	32.4%	Score 1427;	DB 3;	Length 1460;				
	Best Local Similarity	99.3%;	pred. No. 6.4e-295;						
	Matches 1433;	Conservative	0;	Mismatches	10;	Indels	0;	Gaps	0;
Qy	33	AGCCCTGAGGAGCGGACCGATGTGGAAACTGTGCCCCCGCGGGCCCGCAGGAGGAG	92						
Db	5	AGCCCTGAGGAGCGGACCGATGTGGAAACTGTGCCCCCGCGGGCCCGCAGGAGGAG	64						
Qy	93	AACCATACAGACTTTTGACCTGGCGTGTGAGTACGTTGTGCAAGGAAACCTGTGCCATT	152						
Db	65	AACCATACAGACTTTTGACCTGGCGTGTGAGTACGTTGTGCAAGGAAACCTGTGCCATT	124						
Qy	153	TAATTGAAAAATGATCAGTCGATCAGCCGGAATCATGCTGTGTTAACTGCTAACTTTTCTG	212						
Db	125	TGATTGAAATGATCAGTCGATCAGCCGGAATCATGCTGTGTTAACTGCTAACTTTCTG	184						
Qy	213	TAACCAACCTGAGTCAAAACAGATGAATCCTCTGTAATGACATTTAAAGATAATCTTAAGT	272						
Db	185	TAACCAACCTGAGTCAAAACAGATGAATCCTCTGTAATGACATTTAAAGATAATCTTAAGT	244						
Qy	273	ATGGTACCTTTGTTAATGAGGAAATAATGAGAAATGGCTTTTCCCGAATCTTCAAGTCGG	332						
Db	245	ATGGTACCTTTGTTAATGAGGAAATAATGAGAAATGGCTTTTCCCGAATCTTCAAGTCGG	304						
Qy	333	GGGATGGTATTACTTTTGGAGTGTGTTGGAAGTAAATTCAGAAATAGAGATGAGCCTTTGG	392						
Db	305	GGGATGGTATTACTTTTGGAGTGTGTTGGAAGTAAATTCAGAAATAGAGATGAGCCTTTGG	364						
Qy	393	TTGCATGCTCTTCTTGTTTAGATGTCTCTGGGAAAACTGCTTTAAATCAAGCTATATTC	452						
Db	365	TTGCATGCTCTTCTTGTTTAGATGTCTCTGGGAAAACTGCTTTAAATCAAGCTATATTC	424						
Qy	453	AACCTTGGAGATTTACTGTTAAACAAATTTGGACAGAAATGCACCTCCTTGTATGGTAT	512						
Db	425	AACCTTGGAGATTTACTGTTAAACAAATTTGGACAGAAATGCACCTCCTTGTATGGTAT	484						
Qy	513	CAGTGAAGTTACCATTTAAACAAATATGTGCACCTCAATTTGTGACGTCCAATTTGTAAGC	572						
Db	485	CAGTGAAGTTACCATTTAAACAAATATGTGCACCTCAATTTGTGACGTCCAATTTGTAAGC	544						
Qy	573	CAGAAATATTTTACTGAATTCCTGAAAGCAGTTTCAGTCCCAAGACGACCTCCCAAAATTG	632						
Db	545	CAGAAATATTTTACTGAATTCCTGAAAGCAGTTTCAGTCCCAAGACGACCTCCCAAAATTG	604						
Qy	633	AAAGTTTTTACCAACCTCTTGATGAACCATCTATTGGAAGTAAAAATGTTGATCTGTACG	692						
Db	605	AAAGTTTTTACCAACCTCTTGATGAACCATCTATTGGAAGTAAAAATGTTGATCTGTACG	664						
Qy	693	GACGCGAGGAAAGAAACAAATCTTCAAGGGGAAAAACATTTATTTTGTGAATGCCAAAC	752						
Db	665	GACGCGAGGAAAGAAACAAATCTTCAAGGGGAAAAACATTTATTTTGTGAATGCCAAAC	724						
Qy	753	AGCATAAGAAATGAGTTCCGCAAGTTGTCTTTTGGAGGTGGGAGCTAGGTTGATAACAG	812						
Db	725	AGCATAAGAAATGAGTTCCGCAAGTTGTCTTTTGGAGGTGGGAGCTAGGTTGATAACAG	784						
Qy	813	AAGAGAAATGAAGAACAATAATTTCTTTTGGCTCGGGAACTGTTGTTGATACAG	872						
Db	785	AAGAGAAATGAAGAACAATAATTTCTTTTGGCTCGGGAACTGTTGTTGATACAG	844						
Qy	873	GAATAACAACCTCCACAGACCTTAATTTCTGACTGTGCAAGAAATGGATTCAGTCAATAA	932						
Db	845	GAATAACAACCTCCACAGACCTTAATTTCTGACTGTGCAAGAAATGGATTCAGTCAATAA	904						
Qy	933	TGGATATGCTCCAAAGCGCAAGTGCTTTAGACCTATTCTCGAAGCAGAAATGGATTTGGCGG	992						







Qy 397 ATGCTCTTCTGTTTGTAGATGTCCTGGGAAACCTGTTTAAATCAAGCTATATTGCAACT 456  
Db |||||  
Qy 452 TTGTTCTTCTGTTTGTAGATGTCCTGGGAAACCTGTTTAAATCAAGCTATTTTACAGCT 511  
Db |||||  
Qy 457 TGGAGGATTTACTGTTAAACAATTGGACAGAGAAATGCACCTGCTGTCATGTTATCAGT 516  
Db |||||  
Qy 512 TGGAGGACTTTACCGCAACAACTGGACGGAGAAATGTACCACCTTGTCTGCGGAGT 571  
Db |||||  
Qy 517 GAAAGTTTACCATTAAACCAATATGTGCACCTCAATTTGTGGAGCTCCAATTTGTAAGCCAGA 576  
Db |||||  
Qy 572 TAAAGTCACCAATTAAACTATATGTGCACCTCAATTTGTGGAGCTCCAATTTAAACCCAGA 631  
Db |||||  
Qy 577 ATATTTTACTGAATTCCTGAAGCAGTTTCAGTCCAGAGCAGCTCCCAAAATTTGAAG 636  
Db |||||  
Qy 632 ATACTTTCTGATTTCTCAAGACAGTTGAATCTAAGAAACAGCTCCAGATTTGAAG 691  
Db |||||  
Qy 637 TTTTATCCACCTCTTGATGACCACTATTTGGAGTAAATGTTGATCTGTCAGAGC 696  
Db |||||  
Qy 692 TTTTATCCCAACCTTGTATGACCACTATTTGGAGTAAATGTTGATCTGTCAGGGCG 751  
Db |||||  
Qy 697 GCAGGAAGAAACAAATCTTCAAGGAAACAACTTTATATTTTGAATGCCAAACAGCA 756  
Db |||||  
Qy 752 ACATGAAGAAACAGATCTTCAAGGAAACAACTTTGTGTTTAAATGCCAAGCAGCA 811  
Db |||||  
Qy 757 TAAGAAATTCAGTTCCGACGTTGTTTGGAGTGGGAAAGCTAGTTGATTAACAGAGA 816  
Db |||||  
Qy 812 CAAGAACTCAGCTCGGAGTTGTTTCGAGGTGGAGAACCCAGGCTGATGGCAGAAGA 871  
Db |||||  
Qy 817 GAATGAAGAAACATATTTCTTTTGGCTCGGGAACCTGTTGTTGATTAACAGGAAT 876  
Db |||||  
Qy 872 CGAGGAAGGAAACAGAGCTTCTTTTCAGCTCCGGAACCTGCGTTGTTGATGTAGGAAT 931  
Db |||||  
Qy 877 AACAACTCACAGACCTTAATTCCTGACTGTCAAGAAATGGAATTCAGTCAATTAATGGA 936  
Db |||||  
Qy 932 AAGCATACAGCTCATATTTTCACTCCAGAAATGGAATTCATTTGATTAATGGA 991  
Db |||||  
Qy 937 TATGCTCCAAAGCAGGCTTTAGACCTATTCCTGAAGCAGAAATTTGATTTGGCGGTGAT 996  
Db |||||  
Qy 992 TACACTTCAAGGAATGCTCTCAGACCTATTCCTGAAGCGGAGATTTGATTTGCTGTAT 1051  
Db |||||  
Qy 997 TTTTCAGCTACAAAGAAATTTACTGTGATCCTCAGGGCCATCCAGTACAGGATTAAGAC 1056  
Db |||||  
Qy 1052 TTTTATGACTACAGAGAAATTTACTGTAATCCGCGAGGCGCAGCTTTGTACAGAAATTAAGAC 1111  
Db |||||  
Qy 1057 AACAACTCCAGGACCAAGCTTTTCAAGGCGTGTGATTTGATGAAGAACTAATGCAAG 1116  
Db |||||  
Qy 1112 AAGACTCCAGGACCAAGCTTTTCCCAAGTCTGTGAGCCAATGGAATAATATCCCAAG 1171  
Db |||||  
Qy 1117 CGCCCCAGTGAACACTAACATACGTAGCTGACACAGAAATTCAGAGCAAGCAGATACATG 1176  
Db |||||  
Qy 1172 TGCTCCAGTGAATATGACCATACGTAGCTGACACAGAAATTCAGAGCCAGCAGATACATG 1231  
Db |||||  
Qy 1177 ---GGATTTGAGTGAAGGCGCAAGAAATCAAGTCTCCAAATGGAACAAATAATTCAG 1233  
Db |||||  
Qy 1232 TATGCTCTTTGAGTGAAGGACAGAGAAATTAAGATCCCTGGAGTGGAAACAAAGCTCTAG 1291  
Db |||||  
Qy 1234 AATGCTTTTCAAGACGCCACCTGTAAGAGGAGTCTGCAAAACAGCTCTTAATAAA 1293  
Db |||||  
Qy 1292 GAAACTTTTCAAGAAACATTTCAATATAAGGAGGCGCCCTTAACCAAGCTCCAAAGCTAA 1351  
Db |||||  
Qy 1294 TAGTATGATCAATATTTTGGCTAAGATGAGAAATCCCAAACTATCAGCTTTTCAACCAAC 1353  
Db |||||  
Qy 1352 CAACGTAGCATCAGATACGCTGGTTAGAGGAAGACCCCGAGCTATCAGCTTTTCCAAAT 1411  
Db |||||  
Qy 1354 TAAATTTGCCAAGTATAAATAAAGTAAAGATAGGCTTTTTCAGCAGCAGCAGCAACTC 1413  
Db |||||  
Qy 1412 GAAATTTCTGTTGCAATATAAATAAAGGATTTGAGCTTTCTCAGCAGCAGCAG---AACTC 1468  
Db |||||  
Qy 1414 CATCAGAACTACTTTTCAGCGCTTACCAAAAAAGGGAAGGATTCAGAAATCAAGA 1473  
Db |||||  
Qy 1469 CATCAAAAACTACTTCCAGCCATGCACCAAGAAAAAGGGAAGGATTAAGACCAACCCAGA 1528  
Db |||||

Qy 1474 AATGCTTTCATGCAAAATCAGCAAGAATAGAAACGTCCTGTTCTCTTTTGAACAACACACA 1533  
Db |||||  
Qy 1529 GCTGCTCTGTCGCAAAATCATCCAGGATGGAGCTGCTGTTCTCTTTTGAACAACACACA 1588  
Db |||||  
Qy 1534 ACTGCTACACCTCATTTGTTGAAAATATAGGAGCAGCATCTATCTGAGAAATGAGCCTGT 1593  
Db |||||  
Qy 1589 ACTGCGGAGCCCTCACTGTGAAAAGCAAGGAG---CATCAGTCTCAGAAATGCGACCT 1645  
Db |||||  
Qy 1594 GGCACAAACTCAGACAAATAACTTTATTACAGATACAGATTTTAAATCTATTGTGAAAA 1653  
Db |||||  
Qy 1646 GGCAGGGAAGCCGACACCTCATCTGTGGGTGGATGGATAGACTACACAGGAGAG 1705  
Db |||||  
Qy 1654 TTCTGCCAGTAATCTCATGCTGCAGAAAAGCTAAGATCAAAATAAAAAAGGAAAA 1709  
Db |||||  
Qy 1706 TCCTGACAGGAACCACTTCCACAGAAACTCTTAGACCAGAAAAAATAAAAAA 1761  
Db |||||

RESULT 7  
AK031933 1731 bp mRNA linear HTC 03-APR-2004  
LOCUS Mus musculus adult male medulla oblongata cDNA, RIKEN full-length  
DEFINITION enriched library, clone:633050E13 product:nibrin, full insert  
sequence.  
ACCESSION AK031933  
VERSION AK031933.1 GI:26327732  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1  
Carninci, P. and Hayaahizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
10349636  
2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayaahizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159  
3  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kusunai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayaahizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861  
4  
The RIKEN Genome Exploration Research Group Phase II Team and the  
FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)  
5  
The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 1731)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayaahida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kaeukawa, T.,



KatoH,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Onisato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.  
 Direct Submission  
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gs.c.riken.jp, URL:http://genome.gsc.riken.jp/, tel:81-45-503-9222, Fax:81-45-503-9216)  
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
 Please visit our web site for further details.  
 URL:http://genome.gsc.riken.jp/  
 URL:http://fantom.gsc.riken.jp/  
 Location/Qualifiers

## FEATURES

source

1..1731  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="FANTOM DB:6330501B13"  
 /db\_xref="taxon:10090"  
 /clone="6330501B13"  
 /sex="male"  
 /tissue\_type="medulla oblongata"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stage="adult"  
 94..>1731  
 /note="unnamed protein product; nibrin (MGD|MGI:1351625, GBINM\_013752, evidence: BLASTN, 100%, match=1730) putative"  
 /codon\_start=1  
 /protein\_id="BAC27610.1"  
 /db\_xref="GI:26327733"  
 /translation="MWKLPAAGAPGEYVRLLAGVEYVVGKNGILLIENDOSIRN HAVLVNFPVTVLSQDEIPITLIKDNKSKIFVNEEKMQTSLTKTGDRVTFGV ESKRVEYEPFLVSSCLDSVGLVNLQILQGLTANNWTEECETHLWMSAVKVIK TICALFCGRPIIKPEYSEFLKAVESKQPPDI EGFYPPIDEPAIGSKSDVLSGRHER KOIFKGTVPFLNAQKHKLSSAVFGGGEARLMAEDDEEQSFSAFGTCVVDVGIH NTQLIISHQKWIHLIMDTLORNGLRPIPEABIGLAVIFMTTNYCNPOGQPTLEK TTPGPSLSQVLSANGKIIPSAFVNMVTVADTESEPDTCMPLSERPEVVKIPLGEO SRKLSQETFNIEKAPKPSKANNVNSDILVRGTFPSYOLSPMPVANKNDWTSQO QNSIKNYFQPCTRKRERDENPESLSSCKSSRLSELSLEQTQAPGPSLWKSREHQS ONATLDREADTSVSGMDIELNRKSPDRKPLPTETLRP"

## ORIGIN

Query Match 24.6%; Score 1083; DB 3; Length 1731;  
 Best Local Similarity 79.8%; Pred. No. 3.8e-221;  
 Matches 1320; Conservative 0; Mismatches 330; Indels 9; Gaps 3;  
 QY 37 CTGAGAGCGGACCGATGTGAAACTGCTGCCCGCGCGGCGCGCGAGGAGAAACC 96  
 DB 78 CTGAGAGCGGCGCCCATGTGGAAGCTGCTCCGGCGCGCGGTCCGAGGAAACC 137  
 QY 97 ATACAGACTTTTGTGCGCGGTGAGTACGTTGTTGGAAGGAAAAAATCTGTCATTAAT 156  
 DB 138 ATACCGACTTTTGGCGCGGTGAGTACGTTGTTGGGAGGAAAAAATCTGGCATTCGTAT 197  
 QY 157 TGAATATGATCAGTCGATCAGCGGAATCATGCTGTGTTAACTGCTAACTTTCTGTAC 216  
 DB 198 TGAATATGATCAGTCGATCAGTCGAAACCATGCTGCTTAAACAGTAAACTTCCCTGTAA 257  
 QY 217 CAACCTGAGTCAACAGATGAATCCCTGTATTGACATTTAAAGATAAATCTTAAGTATGG 276

DB 258 CAGTTTGTAGTCAAAACAGATGAAATTCCTACATTTAAACAATAAAGATAAATCTTAAGTATGG 317  
 QY 277 TACCTTTGTTAATGAGGAAAAAATCGAGAAATGGCTTTTCCGAACTTTGAAAGTCGGGGA 336  
 DB 318 AACCTTTGTTAATGAGGAAAAAATCGAGATGGTCTTCTCTGACGTTTGAAGACAGGAGA 377  
 QY 337 TGGTATTACTTTTGCAGTGTGTTGGAAGTAAATTTCAAGATAGAGTATGAGCCCTTTGGTTC 396  
 DB 378 TAGAGTTTACCTTTGGGGTGTGTTGAAAGTAAATTCAGAGTAGAATACGAGCCCTTGGTGT 437  
 QY 397 ATGCTCTTCTTTTGTAGATGTCTCTGGGAAAACTCTTTTAAATCAAGCTATATTCAACT 456  
 DB 438 TTGTTCTTCTTGTGTAGATGTCTCTGGGAAAACTGTTTAAATCAAGCTATTTTACAGCT 497  
 QY 457 TGGAGATTTACTGTAAACAATTTGACAGAGAAGATGCACTCACCTTGTCTATGATATCAGT 516  
 DB 498 TGGAGGACTTACCGCAAAACAATGGAACGGAAGAATGTACCCACCTTGTCTGTCTCGGCT 557  
 QY 517 GAAAGTTACCATTTAAACAATATGTGCACTCATTTGTGAGCTGCCAATTTGTAAGCCAGA 576  
 DB 558 TAAAGTCAACATTAACAATATGTGCACTCATTTGTGAGCTCCCAATTTAATAAACAGCA 617  
 QY 577 ATATTTTACTGAATTCCTGAAAGCAGTTTCAAGTCCAAAGAACGAGCTCCCAAAATTTGAAAG 636  
 DB 618 ATACTTTTCTGAATTTCTCAAAGCAGTTTGAATCTAAGAAAACAGCCTCCAGATATTTGAAAG 677  
 QY 637 TTTTACCACCTCTTGTATGAACCATCTATTGGAAGTAAATTTGATCTGTCTCAGGAGC 696  
 DB 678 TTTTACCACCTCTTGTATGAACCATCTATTGGAAGTAAATTTGATCTGTCTCAGGAGC 737  
 QY 697 GCAGAAAAAGAAAAAACAATCTTCAAAGGAAAAACAATTTATTTTGAATGCCAAAAACAGCA 756  
 DB 738 ACATGAAAGAAAAACAGATCTTCAAAGGAAAAACAATTTGTTTAAATGCCAGCAGCA 797  
 QY 757 TAAGAAATTTAGTTCCGAGTTTCTTTGGAGGTGGGAAAGCTAGGTGATTAACAGAGA 816  
 DB 798 CAAGAACTCAGCTCGGAGTTGCTTTCCGAGGTGAGAGAACGAGCTGATGGCAGAGA 857  
 QY 817 GAATGAAGAGAACATAATTTCTTTTGGCTCCGGGAAGCTGTGTTGTTGATACAGGAAT 876  
 DB 858 CGACGAAGAGGAACAGAGATCTTTTTCAGCTCCCGGAACCTGCGTGTGTTGATGTAGGAAT 917  
 QY 877 AACAACTCACAGACCTTAATTCCTGACTGTCTCAGAGAAATAGGATTCAGTCAATATGA 936  
 DB 918 AACGATACAGCTCATTAATTTACATCTCCAGAAAAAATGGATTCATTTGATTAATGA 977  
 QY 937 TATGCTCCAAAGGCAAGGTCTTAGACCTATTCTCTGAAGCAGAAATTTGATTTGGCGGTGAT 996  
 DB 978 TACACTTCAAAGGAATGGTCTCAGACCTATTCTCTGAAGCGGAGATTGGATTGGCTTTAT 1037  
 QY 997 TTTGATGACTACAAGAAATTTACTGTGATCTCAGGGCCATCCAGTACAGGATTAAGAC 1056  
 DB 1038 TTTTATGACTACAGAAATTAATCTGTAATCCGACGCGCCAGCTTGTGACAGAAATTAAGAC 1097  
 QY 1057 AACAACTCCAGGACCAAGCCTTTCAAGGCGGTGTGATGATGAAAACTAATCCCAAG 1116  
 DB 1098 AACGATCCAGGACCAAGCCTTTCCAGTCTCTGTACCAATGGAATAATATCCCAAG 1157  
 QY 1117 GCGCCCCAGTGAACTACAAATACGTAGCTGACACAGAAATCAGAGCAGAGATACATG 1176  
 DB 1158 TGCTCCAGTGAATATGACCAATAGCTAGCTGACACAGAAATCAGAGCAGAGATACATG 1217  
 QY 1177 ---GGAATTTGATGAAAGCCCAAGAAATCAAGTCTCCAAATTTGGACAAAAATTCAG 1233  
 DB 1218 TATGCTTTGATGAAAGCCCAAGAAATCAAGTCTCCCTGAGCTGGAACCAAGCTCTAG 1277  
 QY 1234 AATGCTTTCAAGACGCAACCACTGTAAAGAGTCTTGCAAAAAACAAGCTCTTAATAATAA 1293  
 DB 1278 GAAACTTTCAAGAGAAACATTCATATTAAGAGGAGCCCTTAAACCAAGCTCCAAAGCTAA 1337  
 QY 1294 TAGTATGATCAAAATACATTTTGGCTTAAGATGAGAAATCCCAAACTATCAGCTTTTCAACCAAC 1353  
 DB 1338 CAACGTAGATCAGATACGCTGGTTAGAGGAAAGACCCCGAGCTATCAGCTTTTCTCCAAT 1397







LOCUS BX405939 1021 bp mRNA linear EST 03-MAY-2004  
DEFINITION BX405939 Homo sapiens FETAL LIVER Homo sapiens cDNA Clone  
CS0DM009Y012 3-PRIME, mRNA sequence.

ACCESSION BX405939  
VERSION BX405939.2 GI:46954542  
KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1021)

Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On May 15, 2003 this sequence version replaced gi:30764639.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library

was not normalized. Library was constructed by Life Technologies, a

division of Invitrogen.

This sequence belongs to sequence cluster 4762.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?8=CS0DM009BH06NP1&c=4762.r.

FEATURES

source

Location/Qualifiers

1..1021

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DM009Y012"

/tissue\_type="FETAL LIVER"

/dev\_stage="fetal"

/clone\_lib="Homo sapiens FETAL LIVER"

/notes="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-strand cDNA was digested with Not I and

cloned into the Not I and EcoRV sites of the pCMVSPORT 6

vector. Library was not normalized."

ORIGIN

Query Match 19.2%; Score 844.2; DB 5; Length 1021;

Best Local Similarity 90.1%; Pred. No. 5.5e-170;

Matches 930; Conservative 25; Mismatches 65; Indels 12; Gaps 5;

QY 3290 ACTCTCACCTTCCAAATCTTGTGGAAGTGCATCTACTGCCAGAACCAAAATTAACCTACT 3349

||:||||| 1021 ACKCCACCTCCCAATCTTGTGGAAGTGCATCTACTGCCAGAACCAAACTCACTCACT 962

QY 3350 TCCAACTTCCTGGTCTTGCAGTGGAACTCCAGCTGCCAAGGAGTTAGGAAATGAAGG 3409

||:||||| 961 YCCAACTTCCTGGTCTTGCAGTGGAACTCCAGCTGCCAAGGAGTTAGGAAATGAAGG 902

QY 3410 TCTTTTAAAAGCTTCTCAGCCTCTCAGGGAACAGAAATGGGTGAGCCAACTGCA 3469

||:||||| 901 TC-TCHCTAAAAGCTTCTCAGCCTCTCAGGGAACAGAAATGGGTGAGCCAACTGCA 843

QY 3470 ATTCTTACTACGGCATTGAGACAGTGAATATTGAATATTATAGAGATTATGAAC 3529

||:||||| 842 ATCTTACTACGGCATTGAGACAGTGAATATTGAATATTATAGAGATTATGAAC 783

QY 3530 ACTTAAATTATGATAGTGTATGACATTGATAGAAACATGGGATCTTTAGAGTGAAT 3589

||:||||| 782 ACTTAAATTATGATAGTGTATGACATTGATAGAAACATGGGATCTTTAGAGTGAAT 723

QY 3590 TGACGGGATATTAGTGAATGAATGGAGTCAATTTAGCTCTTTAATAGCCATGATCA 3649

||:||||| 722 TGACGGGATATTAGTGAATGAATGGAGTCAATTTAGCTCTTTAATAGCCATGATCA 663

QY 3650 TAAATACCAAGTGAAGCTGGTGAACATATGGTCTCCATTTTACAGTTAAGGAATATAAT 3709

||:|||||

Db 662 TMAATACCAAGTGAAGCTGGTGAACATATGGTCTCCAATCACAGTTAAGGAACATAAT 603

QY 3710 GGACAGATTATATTGTTTCTCTCATGCCACCAATCCCTTCTTAAGGAAGATCCCTTA 3769

Db 602 GGACAGATCAATATTGTTTCTCTCATGCCACCAATCCCTTCTTAAGGAAGATCCCTTA 543

QY 3770 CTATAGCAGTCTTTTATATTGTCATTTATGAATTAATGAATGAGGAGTCTCTGTACT 3829

Db 542 CNAATAGCAGCCCTCATATTGNCACACTCAGACATAGAAGATGA-GAGCTCTGTACT 484

QY 3830 CCGTCTTTTACAAATATTGGGTGTTGTCAGATTTTCCCTTTTAAACOMTTTCCAAAT 3889

Db 483 CCGTCTTTTACAAATATTGG---TGTGTGAGTCACTCTCCCTCTTAACCATCCCAATC 428

QY 3890 CGGGTGTAGTGGATGTTTCCATTTTGGGTTTAAATTTGATATATCCCTGATAGCTATA 3949

Db 427 GGTGTGATGATG---TCTCACTTGGGTCCTAAATTTGATATATCCCTGATAGCTATA 373

QY 3950 TTGGGTCTATAGAAATCTTTTATACATCTAGATGCAAGTCTCTTCGCGATATAGCTATT 4009

Db 372 TTGGGTCTATAGAAATCTTTTATACATCTAGATGCAAGTCTCTTCGCGATATAGCTATT 313

QY 4010 GAGATATTACACCTAGTCTGTGGTGTACTGTTTCTTTATGCTTTTGTGATGAATAGAAG 4069

Db 312 GAGATATTACACCTAGTCTGTGGTGTACTGTTTCTTTATGCTTTTGTGATGAATAGAAG 253

QY 4070 TTTTAAATTTGACAAAGTCAAAATTTATTTTCTTTTGTGATATTTTCTCTCA 4129

Db 252 TSTTAAATTTGACAAAGTCAAAATTTATTTTCTTTTGTGATATTTTCTCTCA 193

QY 4130 ATTTAAACCCCAAGATTTTCAGATATTCTCTCTATTATATAAACTTTATTTTATATT 4189

Db 132 ATTTAAACCCCAAGATTTTCAGATATTCTCTCTATTATATAAACTTTATTTTATATT 133

QY 4190 GTGATCTACCTTGAATTCATATGATGTTGGAATTTGGAATTTGGAATTTGGAATTTGCC 4249

Db 132 GTGATCTACCTTGAATTCATATGATGTTGGAATTTGGAATTTGGAATTTGGAATTTGCC 73

QY 4250 CCNTACAGTATCCAGTCAATGTAACACTGTTTATTGAAGAATTTATCTTTC-CTCAAT 4308

Db 72 CCNTACAGTATCCAGTCAATGTAACACTGTTTATTGAAGAATTTATCTTTC-CTCAAT 13

QY 4309 AAATTACCTTGC 4320

Db 12 AAATTACCTTGC 1

RESULT 10

BM542698

LOCUS

DEFINITION

5', mRNA sequence.

ACCESSION

BM542698

VERSION

BM542698.1

KEYWORDS

EST.

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

REFERENCE

1 (bases 1 to 1163)

AUTHORS

National Institutes of Health,

Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Cloned through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM12185 row: j column: 12

BM542698 1163 bp mRNA linear EST 20-FEB-2002

AGENCOURT\_6449008 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:5520875

5', mRNA sequence.

BM542698

BM542698.1

GI:18772445

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC http://mgi.nci.nih.gov/.

1 (bases 1 to 1163)

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Cloned through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM12185 row: j column: 12



```
High quality sequence start: 7
High quality sequence stop: 619.
Location/Qualifiers
source
1. .1163
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5520875"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/notes="Organ: uterus, Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 2.1 kb. "
```

```
ORIGIN
Query Match      18.9%; Score 834.2; DB 4; Length 1163;
Best Local Similarity 96.9%; Pred. No. 7.8e-168;
Matches 914; Conservative 0; Mismatches 23; Indels 6; Gaps 6;

QY 686 CTGTCCAGCGCGGAGGAAAGAAACAATCTTCAAGGGGAAACATTTATATATTTTGAAT 745
      |||||
Db 28 CGGTCCAGCGCGGAGGAAAGAAACAATCTTCAAGGGGAAACATTTATATTTTGAAT 87
      |||||

QY 746 GCCAACAGCATGAAGAAATTCAGTTCGCGAGTTGCTTTGGAGGTGGGGAAGCTAGGTTG 805
      |||||
Db 88 GCCAAGCAGCATGAAGAAATTCAGTTCGCGAGTTGCTTTGGAGGTGGGGAAGCTAGGTTG 147
      |||||

QY 806 ATACAGAGAGATGAAGAGACATAATTTCTTTTGGCTCGGGAACTGTGTTGTT 865
      |||||
Db 148 ATACAGAGAGATGAAGAGACATAATTTCTTTTGGCTCGGGAACTGTGTTGTT 207
      |||||

QY 866 GATACAGGAATAACAACCTCACAGACCTTAATTCCTGACTGTGACAGAAATGGATTTCAG 925
      |||||
Db 208 GATACAGGAATAACAACCTCACAGACCTTAATTCCTGACTGTGACAGAAATGGATTTCAG 267
      |||||

QY 926 TCAATATGATATGCTCCAAAGCGAGTCTTAGACCTATTCCTGAAGCAGAAATTTGGA 985
      |||||
Db 268 TCAATATGATATGCTCCAAAGCGAGTCTTAGACCTATTCCTGAAGCAGAAATTTGGA 327
      |||||

QY 986 TTGCGCGTGAATTTTCATGACTACAAAGAAATTAATGATGATCTCAGGCGCCATCCAGTACA 1045
      |||||
Db 328 TTGCGCGTGAATTTTCATGACTACAAAGAAATTAATGATGATCTCAGGCGCCATCCAGTACA 387
      |||||

QY 1046 GGATTAAGACAAACACTCCAGGACCAAGCTTTTCAAGCGCTGTGAGTTGATGAAAA 1105
      |||||
Db 388 GGATTAAGACAAACACTCCAGGACCAAGCTTTTCAAGCGCTGTGAGTTGATGAAAA 447
      |||||

QY 1106 CTAATGCCAAGCGCCCGAGTGAAACATACATACGTAGCTGACACAGAAATCAGAGCAA 1165
      |||||
Db 448 CTAATGCCAAGCGCCCGAGTGAAACATACATACGTAGCTGACACAGAAATCAGAGCAA 507
      |||||

QY 1166 GCAGATACATGGATTTGAGTGAAAGGCCAAAGAAATCAAGTCTCCAAATTTGGAACAA 1225
      |||||
Db 508 GCAGATACATGGATTTGAGTGAAAGGCCAAAGAAATCAAGTCTCCAAATTTGGAACAA 567
      |||||

QY 1226 AAATTCCAGATGCTTTTCAAGACGACCCACTGTAAAGAGTCTTGCAAAACAAAGCTCT 1285
      |||||
Db 568 AAATTCCAGATGCTTTTCAAGATGACCCACTGTAAAGAGTCTTGCAAAACAAAGCTCT 627
      |||||

QY 1286 AATAATAATAGTATGTTATCAAAATATCTTTGGCTTAAGATGAGAAATCCCAACTATCAGC-T 1344
      |||||
Db 628 AATAATAATAGTATGTTATCAAAATATCTTTGGCTTAAGATGAGAAATCCCAACTATCAGC-T 687
      |||||

QY 1345 TTCAACCACTAAATTT-GCCAGATATAAATAAAGTAAAGTAGGGCTTCTCAGCAGCAGC 1403
      |||||
Db 688 TTCAACCACTAAATTTGGCCAGTATAAATAAAGTAAAGTAGGGCTTCTCAGCAGCAGC 747
      |||||

QY 1404 AGACCAACTCCATCAGAACTACTTTTCAGCGCTCTACCAAAAAAGGAAAGGATGAAG 1463
      |||||
Db 748 AGACCAACTCCATCAGAACTACTTTTCAGCGCTCTACCCAAAAGGAAAGGATGAAG 807
      |||||

QY 1464 AAAATCAAGAAATGTCTTTCATGCAAAATCAGCAAGAAATAGAAACGTCTTGTCTCTTTTAG 1523
      |||||
```

```
Db 808 GAAATCCAGATGTCTTTCATTCGAATCAGCAAAATAG-AACGGCTGTCTTCTCTTTTA- 865
      |||||
QY 1524 AACAAACACAACTCTGTACACCTCATTTGTGGAAATAAGGAGCAGCATCTATCTG-AG 1582
      |||||
Db 866 AAAAACAACAACTCTGTACACCTCATTTGTGGAAATAAGGAGCAGCATCTATCTGAAA 925
      |||||

QY 1583 AATGAGCCTGTGGACACAAC-TCAGACAATACTATTATTACA 1624
      |||||
Db 926 AATGAGCCTGTGGAAACCACTTCAGACAATACTATTATTACA 968
      |||||

RESULT 11
BG182890
LOCUS
DEFINITION
RST1768 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION
BG182890
VERSION
BG182890.1
GI:13704577
SOURCE
EST.
KEYWORDS
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 927)
AUTHORS
Harrington, J. J., Sherf, B., Rundlett, S., Jackson, P. D., Perry, R.,
Cain, S., Leventhal, C., Thornton, M., Ramachandran, R.,
Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Booser, S.,
Mays, R., Smith, E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K.,
Offenbacher, J., Danzig, J. and Ducar, M.
Creation of genome-wide protein expression libraries using random
activation of gene expression
Nat. Biotechnol. 19 (5), 440-445 (2001)
JOURNAL
MEDLINE
PUBMED
11329013
COMMENT
Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@athersys.com
High quality sequence stop: 562.
Location/Qualifiers
1. .927
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/notes="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."
```

```
ORIGIN
Query Match      18.7%; Score 822.8; DB 4; Length 927;
Best Local Similarity 94.7%; Pred. No. 2.1e-165;
Matches 859; Conservative 0; Mismatches 47; Indels 1; Gaps 1;

QY 541 TGCACCTCATTTGTGGACGTCCTCAATTTGTAAGCCAGCAATATTTTACTGAATTCCTGAAGC 600
      |||||
Db 21 TGCACCTCATTTGTGGNACGTCCTCAATTTGTAAGCCANATATTTTACTGAATTCCTGAAGC 80
      |||||

QY 601 AGTTTCAGTCCAGAGACGCTCCCAAAATGAAAGTTTTTACCACCTCTTGTGTAAGC 660
      |||||
Db 81 AGTTTCAGTCCAGAGACGCTCCCAAAATGAAAGTTTTTACCACCTCTTGTGTAAGC 140
      |||||

QY 661 ATCTATTGAACTAAATAATGTTGATCTGTCCAGAGCGGAGGAAAGAAACAAATCTTCAA 720
      |||||
Db 141 ATCTATTGAACTAAATAATGTTGATCTGTCCAGAGCGGAGGAAAGAAACAAATCTTCAA 200
      |||||

QY 721 AGGGAAACATTTATATATTTTGAATGCCAAACACAGCATAGAAATTTGAGTTCCGAGTTGT 780
      |||||
```



```
Db 201 AGAGAAAAATTATATTTTGAATCCAAACAGCATAGAATTCAGTTCGCCAGTGT 260
QY 781 CTTTGGAGGTGGGAAGCTAGGTTGATATACAGAGAAGATGAAGAAGACATAATTTCTT 840
Db 261 CTTTGGAGGTGGGAAGCTAGGTTGATATACAGAGAAGATGAAGAAGACATAATTTCTT 320
QY 841 TTTGGCTCCGGGAACCTGTTGTTGATACAGGAATAACAACTCACAGACCTTAATTC 900
Db 321 TTTGGCTCCGGGAACCTGTTGTTGATACAGGAATAACAACTCACAGACCTTAATTC 380
QY 901 TGACTCTCAGAGAAATGGATTTCAGTCAATATGATATGCTCCAAAGGCAAGGCTTTAG 960
Db 381 TGACTCTCAGAGAAATGGATTTCAGTCAATATGATATGCTCCAAAGGCAAGGCTTTAG 440
QY 961 ACCTATTCCTGAAGCAGAAATGGATTTCAGTTCGCGGTGATTTTCATGACTACAAAGATTA 1020
Db 441 ACCTATTCCTGAAGCAGAAATGGATTTCAGTTCGCGGTGATTTTCATGACTACAAAGATTA 500
QY 1021 TGAATCTCAGGGCCATCCAGTACAGGATTAAGACAACAACTCCAGGACCAAGCCTTTC 1080
Db 501 TGAATCTCAGGGCCATCCAGTACAGGATTAAGACAACAACTCCAGGACCAAGCCTTTC 560
QY 1081 ACAAGGCGTGTCTAGTTGATGAAAAAATAATGCCAAGCGCCCAAGTGAACACTACACATA 1140
Db 561 ACAAGGCGTGTCTAGTTGATGAAAAAATAATGCCAAGCGCCCAAGTGAACACTACACATA 620
QY 1141 CGTAGCTGACACAGAAATCAGAGCAGACATACATGGGATTTGAGTGAAGGCCAAAGAA 1200
Db 621 CGTAGCTGACACAGAAATCAGAGCAGACATACATGGGATTTGAGTGAAGGCCAAAGAA 680
QY 1201 AATCAAGTCTCAGAAATCGAACAAAAATTCAGAAATTCAGAAATTCAGAAATTCAGAA 1260
Db 681 AATCAAGTCTCAGAAATCGAACAAAAATTCAGAAATTCAGAAATTCAGAAATTCAGAA 740
QY 1261 AAAGAGTCTCTGCAAAACCAAGCTCTAATAATAGTATGATGATCAATACTTTGGCTAA 1320
Db 741 NAAGAGTCTCTGCAAAACCAAGCTCTAATAATAGTATGATGATCAATACTTTGGCTAA 799
QY 1321 GATGAGAAATCCCAAACTATCAGCTTTTACCAAACTAAATGCGCAAGTATATAAAGTAA 1380
Db 800 GATGAGAAATCCCAAACTATCAGCTTTTACCAAACTAAATGCGCAAGTATATAAAGTAA 859
QY 1381 AGATAGGCTCTCAGCAGCAGACCAACTCCATCAGAACTACTTTTCAGCGCTAC 1440
Db 860 AGATAGGCTCTCAGCAGCAGACCAACTCCATCAGAACTACTTTTCAGCGCTAC 919
QY 1441 CAAAAA 1447
Db 920 CAAAAA 926

RESULT 12
LOCUS AY416036
DEFINITION Mus musculus NBS1 gene, VIRTUAL TRANSCRIPT, partial sequence,
ACCESSION AY416036
VERSION AY416036.1 GI:39771996
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1803)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
```

```
REFERENCE 2 (bases 1 to 1803)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..1803
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1803
/gene="NBS1"
/locus_tag="HCM5756"
ORIGIN
Query Match 18.7%; Score 822.2; DB 9; Length 1803;
Best Local Similarity 71.2%; Pred. No. 3.1e-165;
Matches 988; Conservative 0; Mismatches 393; Indels 6; Gaps 2;
QY 224 AGTCAACAGATGAATCCCTGATTGACATTAAGAGATAATCTAAGTATGTTACCTTT 283
Db 1 AGTCAACAGATGAATCCCTGATTGACATTAAGAGATAATCTAAGTATGTTACCTTT 60
QY 284 GTTAATGAGGAAAAAATCAGAAATCGCTTTTCCCGAACTTTGAACTCGGGGATGTTAT 343
Db 61 GTTAATGAGGAAAAAATCAGAAATCGCTTTTCCCGAACTTTGAACTCGGGGATGTTAT 120
QY 344 ACTTTTGAGTGTGTTGGAAGTAAATTCAGATAGATAGTACGCTTTGTTGTCATGCTCT 403
Db 121 ACTTTTGAGTGTGTTGGAAGTAAATTCAGATAGATAGTACGCTTTGTTGTCATGCTCT 180
QY 404 TCTTGTGTAGATGTCTCTGGGAAAACTGCTTTTAAATCAAGCTATATTCGCAACTGGAG 463
Db 181 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 240
QY 464 TTTACTGTAAACAATTGGACAGAAATGCACTCACTTGTGTATGTTGTTGTTGTTGTT 523
Db 241 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 300
QY 524 ACCATTAAACAAATATGTCATCTGTCAGTCCAACTCCAACTGTTAAAGCCAGAAATTTT 583
Db 301 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 360
QY 584 ACTGAAATCTCGAAAGCAGTTTCAGTCCAAAGAGAGAGCTCCCAAAATGAAAGTTTTC 643
Db 361 TCTGAAATCTCGAAAGCAGTTTCAGTCCAAAGAGAGAGCTCCCAAAATGAAAGTTTTC 420
QY 644 CCACCTTTGATGAACCAATCTATTGGAAGTAAAAATGTTGATCTGTCTAGGACGGAGAA 703
Db 421 CCACCTTTGATGAACCAATCTATTGGAAGTAAAAATGTTGATCTGTCTAGGACGGAGAA 480
QY 704 AGAAAAACAAATCTTCAAGGGGAAAAATTTATTTTGAATGCCAAACAGCATAGAA 763
Db 481 AGAAAAACAAATCTTCAAGGGGAAAAATTTATTTTGAATGCCAAACAGCATAGAA 540
QY 764 TTGAGTTCGCGAGTTGTCTTTGGAGGTGGGAGCTAGTGTTCATACAGAGAGATGAA 823
Db 541 CTCAGCTCGGCAAGTTGCTTTTCGAGGTGGAGAGCCAGGCTGATGGCAGAACGACGAA 600
QY 824 GAAGAACATAATTTCTTTTGGCTCCGGGAAGCTGTTGTTGATACAGGAATAACAAC 883
Db 601 GAGGAACAGAGCTTTCTTTTCAGTCCCGGAACCTCGTGTGTTGATGTTAGGAATAACGAT 660
QY 884 TCACAGACTTAATTTCTGACTGTTCAGAGAAATGATTCAGTCAATATGATATGCTC 943
Db 661 ACACAGCTCATAAATTTTCACACTCCCGAGAAAAATGGATTCATTTGATATGATACAT 720
QY 944 CAAAGGCAAGGTCTTTAGACCTATTCTCTGAAGCAGAAATTTGGATTTGGCGTGTATG 1003
```



Db 721 CAAGGAGTGGTCTCAGACCTATTCTTGAAGCGAGATTGGATTGGCTGTTATTTATG 780  
Qy 1004 ACTACAAAGAAATTAAGTCTCTCAGGCGCATCCAGATGAGGATTAAGACAAACAAT 1063  
Db 781 ACTACAGAGAAATTAAGTCTCTCAGGCGCATCCAGATGAGGATTAAGACAAACAAT 840  
Qy 1064 CAGGAGCAAGCTTTTCAAGGCGGTGTCTAGTGTAGTGAAGAAATTAATGCGAAGGCGCCCA 1123  
Db 841 CCAGGCAAGCTTTTCAAGGCGGTGTCTAGTGTAGTGAAGAAATTAATGCGAAGGCGCCCA 900  
Qy 1124 GTCAACACTACACATAGCTAGCTGACACAGATCAGACGACGACGATGAGGATTTG 1183  
Db 901 GTGAATATGACCATACGTAGCTGACACAGATCAGACGACGACGATGAGGATTTG 960  
Qy 1184 AGTCAAGGCGCAAGAAATTAAGTCTTCAAAATGGAACAAATTAAGTCTTCA 1243  
Db 961 AGTGAAGACCAAGAGAGTAAAGATCCCTGGACTGGAACAAAGCTTAGAACAATTTCA 1020  
Qy 1244 CAAGAGCCACCACTGTAAAGGAGTCTGTCAAAACCAAGCTCTTAATAATAGTATGGTA 1303  
Db 1021 CAAGAAACATTTCAATATAAGGAGGCGCCCTAAACCAAGCTTCAAAAGCTTAACAACGTAGCA 1080  
Qy 1304 TCAATACCTTTGCTTAAGATGAGATCCCAAACTATCAGCTTTTCAACCACTAAATTTGCCA 1363  
Db 1081 TCAGATACGCTGTGTAGAGAAAGACCCCGAGCTATCAGCTTTTCCATGAAATTTCTT 1140  
Qy 1364 AGTATAAATAAAGTAAAGATAGGCTTCTCAGCAGCAGCAGCACTCCATCAGAAAC 1423  
Db 1141 GTTGCAATAAATAAAGTAAAGATAGGCTTCTCAGCAGCAGCAGCACTCCATCAGAAAC 1197  
Qy 1424 TACTTTTCCAGCTGTCTACCAAAAGGAGGATGGAAGAAATCAAGAAATGTCTTTCA 1483  
Db 1198 TACTTCCAGCTGTCTACCAAAAGGAGGATGGAAGAAATCAAGAAATGTCTTTCA 1257  
Qy 1484 TGAATACAGCAGATAGAAAGCTGTCTTTCTTTTGTAGAAACAAACACACACCTGCTTACA 1543  
Db 1258 TGCATAATCATCCAGGATGAGGCTGTCTGTCTTTTGTAGAAACAAACACACCTGCGGA 1317  
Qy 1544 CCCTCATTTGGAAAAATTAAGAGCAGCATCTATCTGAGAAATGAGCTGTGGACACAAAC 1603  
Db 1318 CCTCATCTGGAAGAGCAAGG---AGCATCAGTCTCAGATCGGACCTTGGACAGGGAA 1374  
Qy 1604 TCAGACA 1610  
Db 1375 GCGGACA 1381

RESULT 13  
CN304420  
LOCUS 1700600059211 GRN\_PRENU Homo sapiens cdna 5', mRNA sequence.  
DEFINITION CN304420  
ACCESSION CN304420  
VERSION CN304420.1 GI:47320834  
KEYWORDS EST.  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 785)  
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G. J.,  
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M. S., Mandalam, R.,  
Lebkowski, J. and Stanton, L. W.  
Transcriptome characterization elucidates signaling networks that  
control human ES cell growth and differentiation  
Nat. Biotechnol. 22 (6), 707-716 (2004)  
Contact: Brandenberger R  
Regenerative Medicine  
Geron Corporation  
230 Constitution Drive, Menlo Park, CA 94025, USA  
Tel: 650 473 8658  
Fax: 650 473 7760  
Email: rbrandenberger@geron.com

Insert Length: 785 Std Error: 0.00.  
Location/Qualifiers  
1. .785  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="embryonic stem cell, retinoic acid and  
mitogen-treated hES cell line H7"  
/clone\_lib="GRN PRENU"  
/note="oligo dT primed, full-length enriched cDNA library  
from hES cell line H7 (p29) maintained in feeder-free  
conditions. Embryoid bodies were generated in the presence  
of all-trans retinoic and mitogens."

FEATURES  
source

ORIGIN

Query Match 17.5%; Score 771.4; DB 7; Length 785;  
Best Local Similarity 99.7%; Pred. No. 2.2e-154;  
Matches 783; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 469 TGTAAACAATTTGACAGAGAATGCACTCACCTTGTCTATGATCATAGTGAAGTTTACCAT 528  
Db 1 TGTAAACAATTTGACAGAGAATGCACTCACCTTGTCTATGATCATAGTGAAGTTTACCAT 60  
Qy 529 TAAACAATTTGCACTCATTTGTGGAGCTCCAAATTTGTAAGCCAGATATTTTACTGA 588  
Db 61 TAAACAATTTGCACTCATTTGTGGAGCTCCAAATTTGTAAGCCAGATATTTTACTGA 120  
Qy 589 ATTCTCTGAAGCAGTTTCAGTCCAAAGAGCAGCTCCACAAATTTGAAAGTTTACCACC 648  
Db 121 ATTCTCTGAAGCAGTTTCAGTCCAAAGAGCAGCTCCACAAATTTGAAAGTTTACCACC 180  
Qy 649 TCTTGATGAACCATCTATTGTGAAGTAAATAATTTGTATCTGTCTGAGCAGCAAGAAATGAG 708  
Db 181 TCTTGATGAACCATCTATTGTGAAGTAAATAATTTGTATCTGTCTGAGCAGCAAGAAATGAG 240  
Qy 709 ACAATCTTCAAGGGAACATTTATATTTTGAATGCCAACAGCATAAGAAATGAG 768  
Db 241 ACAATCTTCAAGGGAACATTTATATTTTGAATGCCAACAGCATAAGAAATGAG 300  
Qy 769 TTCCGAGTGTCTTTTGGAGGTGGGAAGCTAGGTGTGATAACAAGAGAAATGAAGAAG 828  
Db 301 TTCCGAGTGTCTTTTGGAGGTGGGAAGCTAGGTGTGATAACAAGAGAAATGAAGAAG 360  
Qy 829 ACATAATTTCTTTTGGCTCCGGGAACGTGTGTGTGTGATACAGGAATTAACAATCTACA 888  
Db 361 ACATAATTTCTTTTGGCTCCGGGAACGTGTGTGTGTGATACAGGAATTAACAATCTACA 420  
Qy 889 GACCTTAATCTCTGACTGTCTAGAGAAATGGATTCAGTCAATATGATATCTCCAAAG 948  
Db 421 GACCTTAATCTCTGACTGTCTAGAGAAATGGATTCAGTCAATATGATATCTCCAAAG 480  
Qy 949 GCAAGTCTTAGACCTATTCTTGAAGCAGAAATTTGGATTGGCGGTGATTTTCATGACTAC 1008  
Db 481 GCAAGTCTTAGACCTATTCTTGAAGCAGAAATTTGGATTGGCGGTGATTTTCATGACTAC 540  
Qy 1009 AAAGAATTTACTGTGATCTCTCAGGCGCATCCAGTACAGATTAAAGACAACTCCAG 1068  
Db 541 AAAGAATTTACTGTGATCTCTCAGGCGCATCCAGTACAGATTAAAGACAACTCCAG 600  
Qy 1069 ACCAAGCTTTTCAAGGCGGTGTCTAGTGAAGAAATTAATGCCAAGCGCCCACTGAA 1128  
Db 601 ACCAAGCTTTTCAAGGCGGTGTCTAGTGAAGAAATTAATGCCAAGCGCCCACTGAA 660  
Qy 1129 CACTTCAACATACGTAGCTGACACAGAAATCAGAGCAAGCAGATACATATGGATTTGAGTGA 1188  
Db 661 CACTTCAACATACGTAGCTGACACAGAAATCAGAGCAAGCAGATACATATGGATTTGAGTGA 720  
Qy 1189 AA-GGCCAAAAGAAATCAAGTCTCCAAATGGAAACAAATTTCAAGATGCTTTTCAAG 1247  
Db 721 AAGGCGAAAAGAAATCAAGTCTCCAAATGGAAACAAATTTCAAGATGCTTTTCAAG 780  
Qy 1248 ACGCA 1252



```

Db          781 ACGCA 785

RESULT 14
CB250418/c
LOCUS
DEFINITION
  784 bp mRNA linear EST 14-FEB-2003
  UI-CF-FN0-age-g-13-0-UI.s1 UI-CF-FN0 Homo sapiens cDNA clone
  UI-CF-FN0-age-g-13-0-UI 3', mRNA sequence.
ACCESSION
CB250418
VERSION
CB250418.1 GI:28390871
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 784)
  Ronaldo,M.F., Lennon,G. and Soares,M.B.
  Normalization and subtraction: two approaches to facilitate gene
  discovery
  Genome Res. 6 (9), 791-806 (1996)
  97044477
  8889548
  Contact: McCray, PB
  McCray Lab
  University of Iowa
  2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
  Tel: 319 356 4866
  Fax: 319 356 7171
  Email: paul-mccray@uiowa.edu
  Tissue Procurement: Dr. M. J. Welsh, University of Iowa
  cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
  cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: Researchers may obtain clones from Research
  Genetics (www.resgen.com) or from Open Biosystems
  (www.openbiosystems.com).
  Seq primer: M13 FORWARD
  POLYA=Yes.

FEATURES
    source
    Location/Qualifiers
    1..784
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="UI-CF-FN0-age-g-13-0-UI"
    /tissue_type="Human Lung Epithelial cells"
    /lab_host="DH10B (Life Technologies) (TI phage resistant)"
    /clone_lib="UI-CF-FN0"
    /notes="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
    modified polylinker; Site 1: EcoR I; Site 2: Not I;
    UI-CF-FN0 is a subtracted cDNA library derived from two
    normalized Human lung epithelial cell libraries (EN1 and
    DU1) The library was subtracted according to according to
    Ronaldo, Lennon and Soares, Genome Research, 6:791-806,
    1996. For additional information, contact:
    bento-soares@uiowa.edu
    TAG_TISSUE=Lung Epithelial Cells Tissue nos 359-368
    TAG_LIB=UI-CF-FN0
    TAG_SEQ=GGCTGTAGGC"

ORIGIN
  Query Match 16.8%; Score 741.4; DB 6; Length 784;
  Best Local Similarity 98.4%; Pred. No. 5.8e-148;
  Matches 748; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1709 ATGGATGATGTGCCATAGAGTGAAGTATTGGAACAGTATTTCAGGACACAAACCA 1768
  |||||||
Db 777 AGGGATGATGTGCCATAGAGTGAAGTATTGGAACAGTATTTCAGGACACAAACCA 718
  |||||||
QY 1769 GAGTTAGAAATTGATGTGAAGTTCAAAACAGGAGGAGATGTCATGTTAGAAAAGG 1828
  |||||||
Db 717 GAGTTAGAAATTGATGTGAAGTTCNAAAACAGGAGGAGATGTCATGTTAGAAAAGG 658
  |||||||
QY 1829 CCAAGGATGGATATAGAAACAAATGACACTTTTCAGTGTGATGAAGCAGTACCAGAAAAGTAGC 1888
  |||||||
Db          657 CCAAGGATGGATATAGAAACAAATGACACTTTTCAGTGTGATGAAGCAGTACCAGAAAAGTAGC 598
  |||||||
QY 1889 AAAATATCTCAAGAAAATGAAATTTGGGAAGAAACCTGAACTCAAGGAAGAGCTCATTATGG 1948
  |||||||
Db 597 AAAATATCTCAAGAAAATGAAATTTGGGAAGAAACCTGAACTCAAGGAAGAGCTCATTATGG 538
  |||||||
QY 1949 TCAGCTTAAGAAATATCTTAACAATGACAAACTTTCAGGATGATAGTGAGATGCTTCCAAA 2008
  |||||||
Db 537 TCAGCTTAAGAAATATCTTAACAATGACAAACTTTCAGGATGATAGTGAGATGCTTCCAAA 478
  |||||||
QY 2009 AAGCTGTTATTGACTGAAATTTAGATCACTGGTGATTTAAAAAATCTTACTTCCAGAAATCCG 2068
  |||||||
Db 477 AAGCTGTTATTGACTGAAATTTAGATCACTGGTGATTTAAAAAATCTTACTTCCAGAAATCCA 418
  |||||||
QY 2069 TCTGCGATAAATGATGATTTATGGTCAACTAAAAAATTTCAAGAAATTTCAAAAAGGTGACA 2128
  |||||||
Db 417 TCTGCGATAAATGATGATTTATGGTCAACTAAAAAATTTCAAGAAATTTCAAAAAGGTGACA 358
  |||||||
QY 2129 TATCCTGGAGCAGGAAAACCTTCCACACATCATTTGGAGGATCAGATCTAATAGCTCATCAT 2188
  |||||||
Db 357 TATCCTGGAGCAGGAAAACCTTCCACACATCATTTGGAGGATCAGATCTAATAGCTCATCAT 298
  |||||||
QY 2189 GCTCGAAAAGAAATACAGAACTAGAGAGTGGCTTAAGCGAGGAATGGAGGTACAAAATCAA 2248
  |||||||
Db 297 GCTCGAAAAGAAATACAGAACTAGAGAGTGGCTTAAGCGAGGAATGGAGGTACAAAATCAA 238
  |||||||
QY 2249 CATGCAAAAGAGAGTCTCTTGCTGATGATCTTTTAGATACATCTTTATTTAAAGG 2308
  |||||||
Db 237 CATGCAAAAGAGAGTCTCTTGCTGATGATCTTTTAGATACATCTTTATTTAAAGG 178
  |||||||
QY 2309 AGAAGTAACCTGAGGATTTTAAAAAGAGCCATGAAAAAATCTCTAGTAGCATCTACT 2368
  |||||||
Db 177 AGAAGTAACCTGAGGATTTTAAAAAGAGCCATGAAAAAATCTCTAGTAGCATCTACT 118
  |||||||
QY 2369 TCAGGCCCAACAGGTTTATATGAATATATAGTGATATAGAGCGAATTAAGTTTACAATGTTT 2428
  |||||||
Db 117 TCAGGCCCAACAGGTTTATATGAATATATAGTGATATAGAGCGAATTAAGTTTACAATGTTT 58
  |||||||
QY 2429 TATGCGCTAAATTTTAAATAAATGCAAAAACTTTGA 2468
  |||||||
Db 57 TATGCGCTAAATTTTAAATAAATGCAAAAACTTTGA 18
  |||||||

RESULT 15
BG214621
LOCUS
DEFINITION
  RST34264 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION
BG214621
VERSION
BG214621.1 GI:13740642
KEYWORDS
  EST.
  Homo sapiens (human)
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
  Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,
  Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,
  Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K.,
  Offenbacher,J., Danzig,J. and Ducar,M.
  Creation of genome-wide protein expression libraries using random
  activation of gene expression
  Nat. Biotechnol. 19 (5), 440-445 (2001)
  21227151
  11329013
  COMMENT
  Contact: Scott J. Cain
  Athersys, Inc.
  3201 Carnegie Ave, Cleveland, OH 44115, USA
  Tel: 216 431 9900
  Fax: 216 361 9596
  Email: scain@athersys.com
  High quality sequence stop: 545.

```



**FEATURES**  
**SOURCE**

Location/Qualifiers  
1. .789

2.102  
/organism="Homo sapiens"

```

/mol type="mrna"

```

```
/db xref="taxon:9606"
```

```

/width= 300
/ceil_line="HT1080"

```

```
/clone lib="Atheros RAGE Library"
```

/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

## ORIGIN

Query Match

Query Match 10.0%, score 741, DS 4, length 183,  
Best Local Similarity 97.8%, Pred. No. 7.1e-148:

Best local similarity	97.8%	Prod. NO. 7.1E-148;
Matches	772; Conservative	0; Mismatches 15; Indels 2; Gaps 2;

Qy	637	TTTTTACCACTCTTGATGAAC-CATCTATTGGAAGTAAAAATGTTGATCTGTCAAGGAC	695
Dβ	2	TTTTTTACCACTCTTGATGAACCTCATCTGGAAGTAAAAATGTTGATCTGTCAAGGAC	61

QY 696 GGCGGGAAGAAAAACAATCTTCAAAGGGAAAAACATTATATTTTGAATGCCAAACAGC 755

Db 62 GGCAGGAAGAAAAACAATCTTCAAGGGAAAAACATTATATTTTGAATGCCAAACAGC 121

Qy 756 ATAAGAAATTGAGTTCGCAAGTTGTCTTTGGAGGTGGGGAAGCTAGTGTGATAACAGAAG 815

Db  
122 ATAGAAATTGAGTTCGAGTTCCTTTGGAGGTGGGGAGGCTAGGTTGATATACAGAAG 181

Qy 816 AGAATGAAGAAGAACATAATTTCTTTTGGCTCCGGGAACGTGTGTTGTGATACAGAA 875

D<sub>b</sub> 182 AGAATGAAGAAGACATAATTTCTTTTGGCTCCGGGAACGTGTGTTGTTGATACAGGAA 241

Qy 876 TAACAACTCACAGACCTTAATCTCTGACTGTGCAGAGAATGGATTGATCAATAATGG 935

Db 242 TAACAAACTCACAGACC-TGATTCTGTGACTGTGAGAAGAAATGGATTTCAGTCAATAATGG 300

[illegible]

DD	301	ATAATGCTCCAAAGGCAAGGCTTACGACCTATCTTGNAGGCAATATAATGTGATTTGGCGGTGA	366
QV	996	TTTTTTCATGACTCAAGAAGATTACTCTGATGATCTCCAGGGCCATCCCACTACAGATTAAAGA	1051

47  
526  
526  
361  
Db

Qy 1056 CAACAACTCCAGGACCAAGCCTTTACAGGCGTGTCAAGTGTGATGAAAACTAATGCCAA 111

Db 421 CAACAACCTCCAGGACCAGGCGCTTTTCAACAGGCGTGTCAAGTGTATGAAAAAATAATGCCAA 480

Qy 1116 GCGCCCCAGTGAACACTACAACATACATCGTAGCTGACACAGAAATCAGAGCAAGCAGATACAT 1117

Db 481 GCGCCCCAGTGAACTACACTACATACATCGTAGCTGACACAGAAATCAGCGCAGCAGATACAT 540

Qy 1176 GGGATTGGAGTGAAGGCCCAAGAAATCAAAGTCTCCAAATGGGAACAAAAATTCAGAA 123

Db 541 GGGATTGTGAGCGAAGGGCCAAAAGAAATCAAAAGTCTCCAAATGGAACAAAAATTCAGAA 600

Qy 1236 TGCTTTCAAGAGCGACCCCACTGTGTAAGGAGTCTTCGCAAAACAAGCTCTTAATAATA 1299

Db 601 TGCTTTCAAGAGCGACCCCACTGTAAAGGAGTCTTGCAAAACAAGCTCTAATAATA 660

1236 GTATGGTATCAATACTTTTGGCTAAGATGAGAATCCCAAACTATCAGCTTTCCACCAACTA 135

DD	661	GCATGGTATCAAAVACCTTTGACTAAGATGAGAAICCCATACATCTTTTACCACCACTA	/21
QV	1355	AATTGCCACGCTATTAATAATAAAGCTAAAAGATACCCGCTTCTACGACGACGACCAACCTCCA	1411

**Vy**

AAATGCGCAAGATATAAATTAATAAAGTAAAAAGTAGGGGCCTTCATCAGCAGCAGCAGACACATCCCA  
1356

**Dβ**

AATTGCCCAAGTATAAATAAAGTAAATGAATGATAGGGCTTCTCATCAGCAGCAGCAGCCAATCTCCA  
721

QY 1416 TCAGAAACT 1424

— — — — —

D<sub>b</sub> 781 TCAGAGACT 789

Search completed: August 30, 2005, 14:45:13  
Job time : 13123 secs



**THIS PAGE BLANK (USPTO)**